

Catherine Feuillet
Gary J. Muehlbauer
Editors

Plant Genetics / Genomics **Volume 7**



Genetics and Genomics of the Triticeae

 Springer

Plant Genetics and Genomics: Crops and Models

Volume 7

Plant Genetics and Genomics: Crops and Models

Series Editor: Richard A. Jorgensen

Forthcoming and planned volumes

- Vol. 1 Genomics of Tropical Crop Plants (eds: Paul Moore/Ray Ming)
- Vol. 2 Genetics and Genomics of Soybean (ed: Gary Stacey)
- Vol. 3 Genetics and Genomics of Cotton (ed: Andy Paterson)
- Vol. 4 Plant Cytogenetics: Genome Structure and Chromosome Function
(eds: Hank Bass/Jim Birchler)
- Vol. 5 Plant Cytogenetics: Methods and Instruction (eds: Hank Bass/Jim
Birchler)
- Vol. 6 Genetics and Genomics of the Rosaceae (eds: Kevin Folta/Sue Gardiner)
- Vol. 7 Genetics and Genomics of the Triticeae (ed: Catherine Feuillet/Gary
Muehlbauer)
- Vol. 8 Genomics of Poplar (ed: Stefan Janssen et al.)

Catherine Feuillet • Gary J. Muehlbauer
Editors

Genetics and Genomics of the *Triticeae*

Foreword by Ronald L. Phillips

 Springer

Editors

Catherine Feuillet
National Institute for Agricultural
Research (INRA)
Genetics, Diversity and Ecophysiology
of Cereals
Clermont-Ferrand 63100
France
catherine.feUILlet@clermont.inra.fr

Gary J. Muehlbauer
Department of Agronomy and Plant
Genetics
University of Minnesota
St. Paul, MN 55108
USA
muehl003@umn.edu

ISBN 978-0-387-77488-6 e-ISBN 978-0-387-77489-3
DOI 10.1007/978-0-387-77489-3
Springer Dordrecht Heidelberg London New York

Library of Congress Control Number: 2009921348

© Springer Science+Business Media, LLC 2009

All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher (Springer Science+Business Media, LLC, 233 Spring Street, New York, NY 10013, USA), except for brief excerpts in connection with reviews or scholarly analysis. Use in connection with any form of information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed is forbidden. The use in this publication of trade names, trademarks, service marks, and similar terms, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Foreword

At this writing, a strain of stem rust of wheat is threatening production around the world, reminiscent of the North American stem rust epiphytotic in 1903 and 1905 and from 1950 to 1954. Norman Borlaug is saying that every commercial wheat variety in the world will need to be replaced with a resistant type. This situation highlights why genetics and genomics information for the Triticeae is so important. Because about half of the yield advances in our major crops generally is the result of genetic advances, any information that can be obtained about genes and gene expression of our rich plant genetic resources is extremely valuable.

Reading this up-to-date book reminds one as much of a textbook as a reference book for learning about plant genetics and systematics as well as the structural and functional organization and evolution of the Triticeae. The book is forward-looking in many ways, including a review of the value of model species to speed genetic understanding of the various crops. The advantages of *Brachypodium* as a model species are described including the fact that it has more traits in common with other Triticeae than other model systems; given that *Arabidopsis* has been a valuable model for all plant species, *Brachypodium* may offer a leap forward for the Triticeae. The large and complex genome of wheat, for example, made initial progress rather slow and made funding agencies wary of investing. This book shows that those days are or at least should be over.

The Triticeae is made up of numerous relatives of approximately 350 species (about 30 genera) among which there are 101 species of 19 genera that are either cultivated or otherwise useful as wild species. Fortunately, the 294 ex situ genebanks in 83 countries holds 1,278,000 accessions. This book provides a well-organized summary of what is held in these genebanks and points out the need for a global database. Two databases, however, are recommended – the Mansfield Database and the USDA online system.

Because the Triticeae has so many relatives, wide hybridization using modern methods for enhancing introgression is especially important. Introgression of specific chromosomes from related species into barley has provided resistance to several diseases. The extensive cytogenetic stocks in the Triticeae have and continue to be important in genomic analysis.

The genomic structure of Triticeae is large and complex, and composed of genes interspersed by huge amounts of repetitive elements. The main repetitive elements in Triticeae are retroelements – principally Copia and Gypsy superfamilies – and can compose 55–70% of the total genomic DNA. Evolution of the genomes reflects a “conservative portion” mainly of genes and single copy DNA, and a “dynamic” portion of transposable elements, duplicated genes and gene fragments. Gene density is shown to be positively correlated with recombination rate. The sequence of *Puccinia graminis*, the causal agent of stem rust, also is now available and even it has 47% of its genome in repetitive elements.

Mapping with Triticeae species is extensively discussed in the book. Although impressive maps exist, it is mentioned that they are not dense enough to foster detailed map-based cloning; interestingly, about half of the genes cloned in the Triticeae (10 out of 19) are disease resistance genes. Relatively few SNP markers are available, and there is a call for further development of such markers. In the book, here are many excellent tips on mapping in terms of appropriate mapping populations, detection of linkage methods, Radiation Hybrid mapping, QTL analysis, and use of multiple environments. An extensive and useful discussion of association mapping provides the needed future directions while reviewing the reported studies and indicating how Association Mapping can compensate for lower marker numbers where whole-genome scans are of low resolution.

Chromosome and chromosome-arm sorting are emphasized in terms of the potential to simplify genome sequencing in the Triticeae. One needs to remember that many chromosome arms in wheat have more DNA than the entire rice genome. Chromosome-arm-specific BAC libraries (the book reports 14 to date) will aid in the study of the structure and evolution of individual chromosomes. The Wheat GeneChip contains one-third of all wheat genes – 55, 052 transcripts. Other new tools available for functional gene analysis in barley and wheat include VIGS (Virus Induced Gene Silencing), RNAi (RNA interference), and TIGS (Transient Induced Gene Silencing). Other useful parts of the book include an extensive discussion on scientific name nomenclature, an excellent accounting of the various species and their origins, and the genetics of many traits especially important in domestication.

Finally, information is presented on a number of quality traits such as B-glucan content, B-amylase, and QTLs for malting quality. Interestingly, the terminal segment of 5HL of barley is reported to carry genes for seed dormancy, pre-harvest sprouting, high malt extract, diastatic power, alpha amylase, and other traits. The genetic analysis of many abiotic stress traits such as drought, cold and heat tolerance, and salinity are reviewed. Also included is a chapter on the genetics of flowering. Experience with flowering shows that the variation in a trait may be due to a sequence far away, such as the Vgt1 QTL in maize that is 70 kb upstream from a flowering effector gene (Ap2-like). Sequencing of Triticeae genomes will allow a detailed scan around such regions of interest and greatly advance the identification of regulatory sequences controlling traits that influence productivity.

Wheaties boxes have used the phrase “Breakfast of Champions”. This book shows that Triticeae genetics is now ready to be the main course. Many genetic resources are now readily available in the Triticeae and molecular genetics tools have been demonstrated to be effective in this set of genomically-complex species. These facts together with the economic value, the ever-present biotic and abiotic stresses, and the greatly increased prices for the Triticeae commodity crops, provide the basis and need for increased investments in research and development.

Ronald L. Phillips
Regents Professor and McKnight Presidential Chair in Genomics
University of Minnesota

Preface

Cereals constitute over 50% of total crop production worldwide (<http://www.fao.org/>) and cereal seeds are one of the most important renewable resources for food, feed, and industrial raw materials. Crop species of the Triticeae tribe that comprises wheat, barley, and rye are essential components of human and domestic animal nutrition. Wheat is grown on 17% of all crop area and represents the staple food for 40% of the world's population. Barley ranks fifth in the world production and is widely used for animal feed and food industry. Rye is second after wheat among grains most commonly used in the production of bread and is also very important for mixed animal feeds and as a source of new alleles for biotic and abiotic stress tolerance in wheat breeding programs. Their domestication in the Fertile Crescent 10,000 years ago ushered in the beginning of agriculture and signified an important breakthrough in the advancement of civilization.

The economic importance of the Triticeae has triggered intense cytogenetic and genetic studies in the past decades that resulted in a breadth of information and tools that have been used to develop wheat, barley and rye varieties with increased yield, improved quality and enhanced biotic and abiotic stress tolerance. In contrast, genomics in the Triticeae lagged behind other plant species, hampered by the size (17 Gb for the bread wheat genome, i.e., 40× the rice genome; 5 Gb for barley and 8 Mb for rye) and complexity (high repeat content, polyploidy) of their genomes. Recently, however, the situation has changed dramatically and the convergence of several technology developments led to the development of a “Genomic toolbox” with new and more efficient resources that supported the establishment of robust genomic programs in the Triticeae. These new capabilities will permit a better understanding of the Triticeae plants biology and support the improvement of agronomically important traits in these essential species.

In this book internationally recognized experts summarize advances of the past decades, synthesise the current state of knowledge of the structure, function, and evolution of the Triticeae genomes and describe progress in the application of this knowledge to the improvement of wheat, barley and rye. The taxonomy, origin and conservation of the huge amount of genetic resources present in the Triticeae families are first reviewed together with basic information about

the genetics and cytogenetics of the three main representatives of the family (wheat, barley and rye). A second section provides the state of the art in the development of tools, resources and methods that have revolutionized our knowledge about the structure, function and evolution of the Triticeae genomes followed by a third section that illustrates the application of these resources to study and improve agronomically important traits such as biotic and abiotic stress resistance, plant development and quality. Finally, three chapters open perspectives into the deployment of new genetic approaches to identify traits and how a better understanding of the organisation of the Triticeae genomes and the ongoing development of new sequencing technologies will support future genome sequencing of these essential crops.

Clermont-Ferrand, France
St. Paul, MN, USA

Catherine Feuillet
Gary J. Muehlbauer

Acknowledgments

The editors want to warmly thank the following persons for their support in the reviewing process of the book chapters

Eduard Akhunov, Kansas State University, USA
Jason Able, University of Adelaide, Australia
François Balfourier, INRA GDEC, France
Michel Bernard, INRA GDEC, France
Mike Bevan, John Innes Center, UK
Gilles Charmet, INRA, GDEC, France
Justin Faris, USDA-ARS, Fargo, USA
Geoffrey Fincher, ACPFG-University of Adelaide, Australia
Perry Gustafson, USDA-ARS, University of Missouri, USA
Jane Rogers, John Innes Center, UK
Peter Langridge, ACPFG-University of Adelaide, Australia
Philippe Leroy, INRA GDEC, France
Peter Morrell, Monsanto Company, St. Louis, MO
Hadi Quesneville, INRA URGI, France
Catherine Ravel, INRA GDEC, France
Oscar Rierra-Lizarazu, Oregon State University, USA
Daryl Somers, Agriculture and Agri-Food Canada-
Winnipeg, Canada
Nabila Yahiaoui, CIRAD, France

Contents

Part I Genetics of the Triticeae

1	Scientific Names in the <i>Triticeae</i>	3
	Mary E. Barkworth and Roland von Bothmer	
2	<i>Triticeae</i> Genetic Resources in <i>ex situ</i> Genebank Collections	31
	Helmut Knüpfper	
3	Domestication of the <i>Triticeae</i> in the Fertile Crescent	81
	Benjamin Kilian, Hakan Özkan, Carlo Pozzi, and Francesco Salamini	
4	Cytogenetic Analysis of Wheat and Rye Genomes	121
	Bikram S. Gill and Bernd Friebe	
5	Applying Cytogenetics and Genomics to Wide Hybridisations in the Genus <i>Hordeum</i>	137
	Andreas Houben and Richard Pickering	
6	Methods for Genetic Analysis in the <i>Triticeae</i>	163
	Abraham Korol, David Mester, Zeev Frenkel, and Yefim Ronin	
7	Genetic Mapping in the <i>Triticeae</i>	201
	Anke Lehmensiek, William Bovill, Peter Wenzl, Peter Langridge, and Rudi Appels	
8	Early Stages of Meiosis in Wheat- and the Role of <i>Ph1</i>	237
	Graham Moore	

Part II Tools, Resources and Approaches

9	A Toolbox for <i>Triticeae</i> Genomics	255
	Etienne Paux and Pierre Sourdille	

10	Chromosome Genomics in the Triticeae	285
	Jaroslav Doležel, Hana Šimková, Marie Kubaláková, Jan Šafář, Pavla Suchánková, Jarmila Číhalíková, Jan Bartoš, and Miroslav Valárik	
11	Physical Mapping in the Triticeae	317
	Nils Stein	
12	Map-Based Cloning of Genes in Triticeae (Wheat and Barley)	337
	Simon Krattinger, Thomas Wicker, and Beat Keller	
13	Functional Validation in the <i>Triticeae</i>	359
	Ingo Hein, Jochen Kumlehn, and Robbie Waugh	
14	Genomics of Transposable Elements in the Triticeae	387
	François Sabot and Alan H. Schulman	
15	Gene and Repetitive Sequence Annotation in the Triticeae	407
	Thomas Wicker and C. Robin Buell	
16	<i>Brachypodium distachyon</i>, a New Model for the Triticeae	427
	John Vogel and Jennifer Bragg	
17	Comparative Genomics in the Triticeae	451
	Catherine Feuillet and Jérôme Salse	
Part III Genetics and Genomics of Triticeae Biology		
18	Genomics of Tolerance to Abiotic Stress in the Triticeae	481
	Marco Maccaferri, Maria Corinna Sanguineti, Silvia Giuliani, and Roberto Tuberosa	
19	Genomics of Biotic Interactions in the Triticeae	559
	Roger P. Wise, Nick Lauter, Les Szabo, and Patrick Schweizer	
20	Developmental and Reproductive Traits in the Triticeae	591
	David A. Laurie	
21	Genomics of Quality Traits	611
	W. Ma, O. Anderson, H. Kuchel, Y. Bonnardeaux, H. Collins, M.K. Morell, P. Langridge, and R. Appels	
Part IV Early messages		
22	Linkage Disequilibrium and Association Mapping in the Triticeae . . .	655
	Mark E. Sorrells and Jianming Yu	

Contents	xv
23 Triticeae Genome Structure and Evolution	685
Jan Dvořák	
24 Wheat and Barley Genome Sequencing	713
Kellye Eversole, Andreas Graner and Nils Stein	
Index	743

Contributors

O. Anderson USDA-ARS, Albany, CA, USA, olin.anderson@ars.usda.gov

R. Appels Centre for Comparative Genomics, Department of Agriculture, Murdoch University, rappels@cgc.murdoch.edu.au

Mary E. Barkworth Department of Biology, Utah State University, Salt Lake City, UT, USA, mary@biology.usu.edu

Jan Bartoš Laboratory of Molecular Cytogenetics and Cytometry, Institute of Experimental Botany, Sokolovská 6, CZ-77200 Olomouc, Czech Republic, bartos@ueb.cas.cz

Y. Bonnardeaux CSIRO Mining and Exploration, Perth, WA 6151, Australia, yumiko@cbbc.murdoch.edu.au

Roland von Bothmer Department of Plant Breeding and Biotechnology, Swedish Agricultural University, Umeå, Sweden, Roland.von.Bothmer@ltj.slu.se

William Bovill Centre for Systems Biology, University of Southern Queensland, Toowoomba, Australia; School of Food, Agriculture and Wine, University of Adelaide, Adelaide, Australia, william.bovill@adelaide.edu.au

Jennifer Bragg USDA-ARS Western Regional Research Center, Genomics and Gene Discovery Unit, Albany, CA 94710, USA, jbragg@pw.usda.gov

C. Robin Buell Department of Plant Biology, Michigan State University, East Lansing MI 48824, USA, buell@msu.edu

Jarmila Číhalíková Laboratory of Molecular Cytogenetics and Cytometry, Institute of Experimental Botany, Sokolovská 6, CZ-77200 Olomouc, Czech Republic; Department of Cell Biology and Genetics, Palacký University, Šlechtitelů 11, CZ-78371 Olomouc, Czech Republic, cihalikova@ueb.cas.cz

H. Collins Australian Centre for Plant Functional Genomics, Adelaide, SA 50000, Australia, helen.collins@acpfg.com.au

Jaroslav Doležel Laboratory of Molecular Cytogenetics and Cytometry, Institute of Experimental Botany, Sokolovská 6, CZ-77200 Olomouc, Czech

Republic; Department of Cell Biology and Genetics, Palacký University, Šlechtitelů 11, CZ-78371 Olomouc, Czech Republic, dolezel@ueb.cas.cz

Jan Dvořák Department of Plant Sciences, University of California, Davis, Oakland, CA, USA, jdvorak@ucdavis.edu

Kellye Eversole Eversole Associates, Bethesda, MD, USA, eversole@eversoleassociates.com

Catherine Feuillet INRA-UBP UMR 1095, Genetics, Diversity and Ecophysiology of Cereals, Domaine de Crouelle, 234 Avenue du Brezet, F-63100 Clermont-Ferrand, France, catherine.feUILlet@clermont.inra.fr

Zeev Frenkel Institute of Evolution, Faculty of Science, University of Haifa, Haifa, Israel, frenkel@research.haifa.ac.il

Bernd Friebe Department of Plant Pathology, Kansas State University, Manhattan KS 66506, USA, friebe@ksu.edu

Bikram S. Gill Department of Plant Pathology, Kansas State University, Manhattan KS 66506, USA, bsgill@ksu.edu

Silvia Giuliani Department of Agroenvironmental Sciences and Technology, University of Bologna, Bologna, Italy, silgiul27@yahoo.com

Andreas Graner Leibniz-Institute of Plant Genetics and Crop Plant Research (IPK), Genebank, graner@ipk-gatersleben.de

Ingo Hein Scottish Crop Research Institute, Invergowrie, Dundee DD2 5DA, UK, ingo.hein@sri.ac.uk

Andreas Houben Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Corrensstr. 3, 06466 Gatersleben, Germany, houben@ipk-gatersleben.de

Beat Keller Institute of Plant Biology, University of Zürich, Zollikerstr 107, CH 8008 Zürich, Switzerland, bkeller@botinst.unizh.ch

Benjamin Kilian Institute of Botany III, Heinrich-Heine-Universität Düsseldorf, Universitätsstrasse 1, 40225 Düsseldorf, Germany; Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Genebank/Genome Diversity, Corrensstrasse 3, 06466 Gatersleben, Germany; Max Planck Institute for Plant Breeding Research, Department of Plant Breeding and Genetics, Carl-von-Linné-Weg 10, 50829 Köln, Germany, kilian@ipk-gatersleben.de

Helmut Knüpffer Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Genebank Department, Corrensstrasse 3, D-06466 Gatersleben, Germany, knupffer@ipk-gatersleben.de

Abraham Korol Institute of Evolution, Faculty of Science, University of Haifa, Haifa, Israel, korol@research.haifa.ac.il

Simon Krattinger Institute of Plant Biology, University of Zürich, Zollikerstr 107, CH 8008 Zürich, Switzerland, skratt@botinst.uzh.ch

Marie Kubaláková Laboratory of Molecular Cytogenetics and Cytometry, Institute of Experimental Botany, Sokolovská 6, CZ-77200 Olomouc, Czech Republic; Department of Cell Biology and Genetics, Palacký University, Šlechtitelů 11, CZ-78371 Olomouc, Czech Republic, kubalakova@ueb.cas.cz

H. Kuchel Australian Grain Technologies, Adelaide, SA 50000, Australia, haydn.kuchel@ausgraintech.com

Jochen Kumlehn Leibniz Institute of Plant Genetics and Crop Plant Research, Corrensstrasse 3, 06466 Gatersleben, Germany, kumlehn@ipk-gatersleben.de

P. Langridge Australian Centre for Plant Functional Genomics, University of Adelaide, peter.langridge@acpfg.com.au

David A. Laurie Department of Crop Genetics, John Innes Centre, Norwich Research Park, Norwich, NR4 7UH, UK, david.laurie@bbsrc.ac.uk

Nick Lauter Corn Insects and Crop Genetics Research, USDA-ARS, Department of Plant Pathology and Center for Plant Responses to Environmental Stresses, Iowa State University, Ames, IA 50011-1020, USA, nickl@iastate.edu

Anke Lehmensiek Centre for Systems Biology, University of Southern Queensland, Toowoomba, Australia, lehmensie@usq.edu.au

W. Ma Department of Agriculture and Food, Perth, WA 6151, Australia, w.ma@murdoch.edu.au

Marco Maccaferri Department of Agroenvironmental Sciences and Technology, University of Bologna, Bologna, Italy, marco.maccaferri@unibo.it

David Mester Institute of Evolution, Faculty of Science, University of Haifa, Haifa, Israel, demseter@research.haifa.ac.il

Graham Moore John Innes Centre, Colney, Norwich. NR4 7UH, UK, graham.moore@bbsrc.ac.uk

M.K. Morell CSIRO Food Futures Flagship, Canberra, Australia, matthew.morell@csiro.au

Hakan Özkan Faculty of Agriculture, Department of Field Crops, University of Cukurova, 01330 Adana, Turkey, hozkan@cukurova.edu.tr

Etienne Paux INRA UMR1095 Génétique, Diversité and Ecophysiologie des Céréales, Domaine de Crouël, 234 Avenue du Brézet, 63100 Clermont-Ferrand France, etienne.paux@clermont.inra.fr

Richard Pickering New Zealand Institute for Plant and Food Research Limited, Private Bag 4704, Christchurch 8140, New Zealand, pickeringr@crop.cri.nz

Carlo Pozzi Fondazione Parco Tecnologico Padano, Via Einstein – Localita Cascina Codazza, 26900 Lodi, Italy, carlo.pozzi@unimi.it

Yefim Ronin Institute of Evolution, Faculty of Science, University of Haifa, Haifa, Israel, efim@research.haifa.ac.il

François Sabot MTT/BI Plant Genomics Laboratory, Institute of Biotechnology, University of Helsinki, Helsinki, Finland, francois.sabot@gmail.com

Jan Šafář Laboratory of Molecular Cytogenetics and Cytometry, Institute of Experimental Botany, Sokolovská 6, CZ-77200 Olomouc, Czech Republic, safar@ueb.cas.cz

Francesco Salamini Fondazione Parco Tecnologico Padano, Via Einstein – Localita Cascina Codazza, 26900 Lodi, Italy; Max Planck Institute for Plant Breeding Research, Department of Plant Breeding and Genetics, Carl-von-Linné-Weg 10, 50829 Köln, Germany, francesco.salamini@tecnoparco.org

Jérôme Salse INRA-UBP UMR 1095, Domaine de Crouelle, 234 Avenue du Brezet, F-63100 Clermont-Ferrand, Genetics, Diversity and Ecophysiology of Cereals, France, jsalse@clermont.inra.fr

Maria Corinna Sanguineti Department of Agroenvironmental Sciences and Technology, University of Bologna, Bologna, Italy, maria.sanguineti@unibo.it

Alan H. Schulman MTT/BI Plant Genomics Laboratory, Institute of Biotechnology, University of Helsinki, Helsinki, Finland; MTT Agrifood Research Finland, Biotechnology and Food Research, Plant Genomics, Jokioinen, Finland, alan.schulman@helsinki.fi

Patrick Schweizer Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Gatersleben, Corrensstrasse 3, 06466 Gatersleben, Germany, schweiz@ipk-gatersleben.de

Hana Šimková Laboratory of Molecular Cytogenetics and Cytometry, Institute of Experimental Botany, Sokolovská 6, CZ-77200 Olomouc, Czech Republic; Department of Cell Biology and Genetics, Palacký University, Šlechtitelů 11, CZ-78371 Olomouc, Czech Republic, simkovah@ueb.cas.cz

Mark E. Sorrells Department of Plant Breeding and Genetics, Cornell University, Ithaca, NY, USA, mes12@cornell.edu

Pierre Sourdille INRA UMR1095 Génétique, Diversité and Ecophysiologie des Céréales, Domaine de Crouël, 234 Avenue du Brézet, 63100 Clermont-Ferrand France, pierre.sourdille@clermont.inra.fr

Nils Stein Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Genebank Department, Gatersleben, Germany, stein@ipk-gatersleben.de

Pavla Suchánková Laboratory of Molecular Cytogenetics and Cytometry, Institute of Experimental Botany, Sokolovská 6, CZ-77200 Olomouc, Czech Republic, suchankova@ueb.cas.cz

Les Szabo Cereal Disease Laboratory, USDA-ARS, University of Minnesota, St. Paul, MN 55108, USA, lszabo@umn.edu

Roberto Tuberosa Department of Agroenvironmental Sciences and Technology, University of Bologna, Bologna, Italy, roberto.tuberosa@unibo.it

Miroslav Valárik Laboratory of Molecular Cytogenetics and Cytometry, Institute of Experimental Botany, Sokolovská 6, CZ-77200 Olomouc, Czech Republic, valarik@ueb.cas.cz

John Vogel USDA-ARS Western Regional Research Center, Genomics and Gene Discovery Unit, Albany, CA 94710, USA, john.vogel@ars.usda.gov

Robbie Waugh Scottish Crop Research Institute, Invergowrie, Dundee DD2 5DA, UK, rwaugh@scri.sari.ac.uk

Peter Wenzl Triticarte Pty Ltd and Diversity Arrays Technology Pty Ltd, Canberra, Australia, p.wenzl@DiversityArrays.com

Thomas Wicker Institute of Plant Biology, University of Zurich, Zollikerstrasse 107, CH-8008 Zurich, Switzerland, wicker@botinst.uzh.ch

Roger P. Wise Corn Insects and Crop Genetics Research, USDA-ARS, Department of Plant Pathology and Center for Plant Responses to Environmental Stresses, Iowa State University, Ames, IA 50011-1020, USA, rpwise@iastate.edu

Jianming Yu Department of Agronomy, Kansas State University, Manhattan KS 66506, USA, jyu@ksu.edu

Chapter 1

Scientific Names in the *Triticeae*

Mary E. Barkworth and Roland von Bothmer

Abstract The grass tribe *Triticeae* has been the focus of many research programs because its inclusion of wheat, barley, and rye makes it of critical importance to the world's food supply, an importance that is enhanced by the many other species that are important for forage and soil stabilization. One consequence of the tribe's importance is that scientists throughout the world are engaged in its study, particularly its cultivated species. The crop species are also used as model organisms in research. This is leading to a rapid accumulation of knowledge about the cultivated species and their close relatives and a slower accumulation of knowledge about the other species. For this reason, and because the tribe grows in almost all temperate regions of the world, many different taxonomic treatments have been proposed for its members. As a result, many of its members have more than one correct scientific name and some names have multiple interpretations. Examples are provided of how such situations arise. This is followed by a discussion of the criteria used in selecting a treatment to be used, brief characterizations of the generic interpretations adopted, and summaries of some alternative interpretations.

1.1 The *Triticeae*

Triticeae is the scientific name for the tribe (group) of grasses that includes the cereals *Triticum aestivum* (bread wheat), *T. durum* [\equiv *T. turgidum* ssp. *durum*] (*durum* wheat), *Secale cereale* (rye), and *Hordeum vulgare* (barley), the modern cereal \times *Triticosecale* (triticale), plus about 350 other species (Löve 1984) [Authors for all names mentioned in this chapter are given in the Appendix]. The cereals are undoubtedly the tribe's best known members, but many of the other species are important for forage and soil stabilization. The cereal and annual species are most abundant in western Asia and around the Mediterranean, but the

M.E. Barkworth (✉)

Department of Biology, Utah State University, Salt Lake City, UT, USA

e-mail: mary@biology.usu.edu

other species are found in almost all temperate regions of the world, the country with the greatest concentration being China (Barkworth et al. 2005).

Morphologically, the *Triticeae* differ from other grasses in their combination of open leaf sheaths, membranous ligules, sessile to almost sessile spikelets, and ovaries with a hairy top. Within the tribe, as in other grass tribes, there is variation in many respects including morphology, life cycle, reproductive behavior, ploidy level, genomic constitution, habitat preference, and phenotypic plasticity. Where the *Triticeae* differ from other tribes are in their economic importance and the extent to which they hybridize, forming at least partially fertile hybrids. Because of their economic importance, the *Triticeae* are the focus of research studies by scientists in diverse disciplines and many different countries; because of the rather low barriers to hybridization, almost all members of the tribe are potential resources for development of genetically improved strains and cultivars (e.g., Verushkine and Shechurdine 1933; Ananthawat-Jonsson et al. 1997; Pinto et al. 2003). Even so, most studies have focused on the cereal species or their close relatives (e.g., Dhaliwal et al. 1990), with the result that there is more information available about them than about other members of the tribe.

1.2 Why so Many Names?

A taxonomic treatment is an attempt to partition the observed variation in a group of organisms into a hierarchical set of units or *taxa* [singular form, *taxon*]. Within the *Triticeae*, the commonly used ranks are, in descending order: genus, species, subspecies, and variety. Early taxonomists relied solely on morphological characters in deciding how to partition the tribe into taxa. Today, although people (including taxonomists) like to be able to distinguish taxa morphologically, information from other disciplines is used in developing taxonomic treatments. As more information becomes available, taxonomic treatments need to change to reflect this information, but how, when, and to what extent such changes should be made is often a matter of dispute. Consequently, there are often multiple taxonomic treatments for a group in use at any given time. This is particularly true of widely distributed and well-studied groups such as the *Triticeae*.

The existence of multiple taxonomic treatments means that many taxa have multiple correct scientific names—and that some scientific names have multiple meanings—because the correct scientific name for a taxon depends on the taxonomic treatment adopted. Most taxonomic treatments of the *Triticeae* are incomplete. Many are regional in scope (e.g. Tsvelev 1976; Edgar and Connor 2000; Wu et al. 2006; Wu and Raven 2007; Barkworth et al. 2007); others treat only one genus (e.g., Dorofeev et al. 1979; Frederiksen 1986, 1991a,b; Bothmer et al. 1995; Baden 1997; Baden et al. 1997; Frederiksen and Peterson 1998). A few discuss the genera to be recognized but not the

delimitation of species in each genus (Dewey 1984; Yen et al. 2005a). Löve's (1984) treatment of the *Triticeae* is the only one that treats all members of the tribe. His work is incredibly valuable and shows considerable insight. Nevertheless, it contains aspects which all, or almost all, of those working with the tribe now reject. This is, however, true of most taxonomic treatments—and taxonomists themselves will change their treatment of a group as new information is obtained or because of changing generic or species concepts (cf. *Elytrigia* in Dewey 1982, 1984; *Critesion* in Barkworth and Dewey 1985; Barkworth et al. 2007).

There are no rules for making taxonomic decisions. The primary goal of the International Code for Botanical Nomenclature (henceforth, "Code"; McNeill et al. 2006) is to ensure that there is one, and only one, correct name for each plant. It provides rules, called Articles, for creating names or deciding among existing names for a taxon, but these come into play only after a taxon's circumscription (i.e., its taxonomic treatment) has been decided.

1.2.1 Impact of New Technologies on the Taxonomy of the Triticeae

The current taxonomic tradition, in which an attempt is made to place related organisms together, dates from around Theophrastus' time (Pavord 2005). The use of binomials for naming species was introduced by Bauhin in 1650 (see Lawrence 1968), but the first *International Code of Botanical Nomenclature* was not published until 1930 although there had been several earlier attempts to reach agreement (see Lawrence 1968). Early taxonomists placed plants that were evidently distinct in different species; genera comprised species that resembled each other more than species of another genus. Thus Bentham (1882), Hackel (1887), and Hitchcock (1951) interpreted *Agropyron* as including all the species of the tribe with solitary spikelets. As knowledge accumulated, it became evident that some of the evidently distinct genera included a mix of closely and distantly related species which, as understanding of evolution developed, was seen as undesirable, and led to the preparation of revised taxonomic treatments.

Nevski (1934) was the first to attempt realigning the genera of the *Triticeae* so that they more closely reflected the evolution of the tribe. At that time, the impact of the ability to stain chromosomes was reflected in such works as Avdulov's (1931) caryosystematic study of grasses. Nevski was aware of Avdulov's work but he noted that he had to make his generic decisions while there was still only minimal cytological data available for the tribe.

In 1984, the accumulated cytogenetic data led Löve and Dewey to propose, independently of each other, that generic circumscription in the tribe be based on genomic constitution. Each noted, however, that as additional species were studied it might become necessary to recognize more genera. These ideas were

greeted with considerable criticism at the time (Barkworth 2000; Barkworth and Bothmer 2005 for review), but many of circumscriptions accepted today reflect the impact of cytogenetic information.

Since 1984, molecular tools have become the “new technology”. They have been employed in a wide range of studies of the *Triticeae*, some focusing on elucidation the relationships among the tribe’s members, others on obtaining information that will aid plant breeders. Integrating such information into the taxonomy, and hence the names used, is still in progress. It has confirmed many, but not all, of the conclusions reached on the basis of cytogenetic information. In some instances, it has made it easier to understand why some taxonomic decisions are problematic. For instance, some molecular studies support treatment of *Triticum* and *Aegilops* as a single genus; others suggest that they should be separate but that *Amblyopyrum* should be included in *Triticum*; others suggest that, if genera are to be monophyletic, the tribe should be reduced to a very small number of genera (see Petersen et al. 2006 for a review).

Once again, a new technology is providing greater insight into the evolution and phylogeny of the tribe. Like other technologies, it is helping us better understand the complexity of the evolutionary and phylogenetic processes that are reflected in the diversity of the tribe. Integrating the implications of such information into the taxonomic treatment of the tribe is an ongoing process.

1.2.2 Integrating New Information into the Taxonomy of the Triticeae

Taxonomists vary in the rate with which they incorporate new information into their treatments. There are many reasons for this, such as differences in the rate at which information comes to the attention of taxonomists and the importance they attach to different kinds of information. Some taxonomists attach considerable importance to having morphologically distinct entities; others think reflecting genetic similarity more important; others consider molecular phylogenetic information paramount. At the generic level, some consider that supraspecific groups such as genera must be monophyletic, that is, have a single common ancestor, a requirement that would reduce the perennial *Triticeae*, if not the whole tribe, to a single genus. Others (e.g., Löve 1984; Dewey 1984; Yen et al. 2005a) emphasize the importance of genomic constitution, seeing this as a measure of genetic relatedness. This is only feasible in the very few groups, the *Triticeae* being the prime example, for which this information is available. Still others emphasize stability. This last is an understandable desire, but one that can require ignoring the implications of new knowledge. All these variables affect the extent to which a taxonomic treatment reflects existing information.

In the remainder of this chapter, we first illustrate why some taxa have multiple names and some names have multiple meanings. We then outline the

guidelines that we followed in deciding on the taxonomic treatment (and hence the names) we currently recommend for use. We conclude with a brief characterization of the genera that we recognize and a summary of alternative treatments. The index includes all the names used in this volume, plus some of their frequently encountered synonyms.

1.3 Interaction of Taxonomy and Nomenclature—Some Examples

In this section, we illustrate why taxonomic treatments change, and the impact of such changes on names, by considering three examples from the *Triticeae*. In the process, we discuss a few of the intricacies of the Code and the reason why it is sometimes important to cite the authors of a scientific name.

1.3.1 Multiple Names at the Generic Level: *Pseudoroegneria*

The impact of different generic treatments on the names of species and lower ranks is easier to follow than differences in the treatment of species and lower ranks. In most instances, all that is involved is replacement of the generic name and, possibly, a change to the ending of the specific epithet so that its gender agrees with that of the generic name. There is often no choice in the specific epithet because the Code (McNeill et al. 2006, Art. 11.4) requires that one use the first specific epithet available. This is the Principle of Priority. Thus, the western North American species now known as *Pseudoroegneria spicata* was originally called *Festuca spicata* by Pursh (1813). In *Agropyron*, it becomes *Agropyron spicatum* (Scribner and Smith 1897), the ending of the epithet changing because *Agropyron* is neuter whereas *Festuca* is feminine. Other nomenclaturally correct names for the taxon are *Elymus spicatus*, published by Gould (1947) and *Elytrigia spicata*, published by Dewey (1983). Which name is used depends on the taxonomic treatment adopted for the genera in the tribe (see Table 1.2). If authors of names are being cited, the citations for the combinations mentioned above are: *Festuca spicata* Pursh, *Agropyron spicatum* (Pursh) Scribn. and J.G. Sm., *Elymus spicatus* (Pursh) Gould, *Pseudoroegneria spicata* (Pursh) Á. Löve, and *Elytrigia spicata* (Pursh) D.R. Dewey. The decision to include both awned and unawned plants in *Pseudoroegneria spicata* was based on ecological, and experimental studies (Daubenmire 1939, 1960; Carlson 2007).

Note that the name of a species is a binomial; the word(s) in roman type that follows the name of a species refer to the author(s) of the name. The Code (McNeill et al. 2006, Art. 46) states that “In publications, particularly those dealing with taxonomy and nomenclature, it *may* be desirable . . . to cite the author(s) of the name concerned” [Emphasis added]. In practice, most editors require citation of the author(s) for all names used at the species level; some

require it for higher ranks. When doing so, it is best to follow the recommendations of Brummitt and Powell (1992) on how, and whether, to abbreviate them. Their recommendations are also available at <http://www.ipni.org>. The Web site also enables one to find the accepted abbreviation for one that is no longer in use. For instance, searching for the abbreviation “Linn.” will bring up the name Linnaeus and its accepted abbreviation, “L.”.

1.3.2 Multiple Names at the Generic Level: Elymus

In some cases, it is impossible to retain the original specific epithet when moving a species to a different genus. For instance, when the western North America species *Agropyron dasystachyum* is included in *Elymus*, it has to be called *Elymus lanceolatus* because, when Gould went to transfer the species in 1949, he found that Trinius (1829) had already used the combination *Elymus dasystachys* for a central Asian species, one now known as *Leymus secalinus* (Wu et al. 2006). When he made the transfer, Gould had three names to consider: *Agropyron dasystachyum*, *A. lanceolatum* and *A. subvillosum*. The authors of these names believed that they referred to three different species, but Gould thought that all three names referred to variants of a single species. Of the three epithets, “*dasystachyum*” was unavailable because it had been used by Trinius. Of the other two epithets, “*lanceolatum*” was first used at the species level by Scribner and Smith in 1897, “*subvillosum*” by Nelson in 1904. Thus, “*lanceolatus*” was the earliest available epithet at the species level so, in accordance with the Principle of Priority, *Agropyron dasystachyum* became *Elymus lanceolatus*, the author citation being (Scribn. & J.G. Sm.) Gould.

1.3.3 Additional Problems with Generic Changes

Determining the correct combination to use when transferring a species from one genus to another is only one of the problems associated with changing generic boundaries. Another is that one must also change descriptions and identification keys. Descriptions of *Agropyron* that reflect its current interpretation usually refer to its closely spaced spikelets with keeled glumes, characteristics that were not true of the majority of the species included in the genus by Bentham (1882). Similarly, the expansion of *Elymus* to include many, but not all, species that are now excluded from *Agropyron* has led to changes in its generic description.

Generic changes also require that care must be taken in interpreting such statements as “*Elymus* is the most widespread genus in the tribe, being native in both the northern and southern hemispheres”. This is true if one accepts the interpretation of Löve (1984) or Edgar and Connor (2000), even though Edgar and Connor interpreted the genus somewhat differently from Löve. It is not

true if one adopts the interpretation of Yen et al. (2005a) because these authors place the Australasian species into other genera.

At a practical level, a greater problem is presented when a generic change is proposed but, for one reason or another, new names are published for only some of its species. For instance, Church (1967) demonstrated that *Hystrix patula*, the type species of *Hystrix* is genetically close to species of *Elymus*. For this reason, it is now included in *Elymus* and is known by the name first given to it by Linnaeus (1753), *E. hystrix*. This does not mean that all the other species that used to be included in *Hystrix* belong in *Elymus*. Indeed, Zhang and Dvorák (1991) and Jensen and Wang (1997) demonstrated that the species known as *Hystrix californica* is genetically more closely related to species of *Leymus* than species of *Elymus*; it was transferred to *Leymus* by Barkworth in 2006. The problem is that most species of *Hystrix* grow in eastern Asia. The authors of the *Flora of China* and Bothmer et al. (2005) decided to recognize *Hystrix* as a genus, consequently there are no names for most of the Chinese and Japanese species in *Elymus* or *Leymus*. Moreover, because few of the species of *Hystrix* have been examined by those who consider that the genus should not be recognized, there is no way of knowing where the species should be placed. All that is clear is that, according to the Code, if *Elymus hystrix*, the type species of *Hystrix*, is included in *Elymus*, *Hystrix* cannot be used as a generic name. A similar problem exists with respect to *Elytrigia*. The matter can only be addressed by studying the species involved.

1.3.4 Multiple Names at the Species Level and Below: The *Triticum monococcum* Complex

Recent changes in the taxonomic treatment of the *Triticum monococcum* complex will illustrate their impact on the names used and the meaning of those names. The focus is on plants associated with five epithets, “*monococcum*”, “*aegilopoides*”, “*boeoticum*”, “*thaoudar*”, and “*urartu*”. Each of these epithets refers to a group of plants that at least one taxonomist has considered worthy of recognition as a taxon.

Four of the five taxa have nomenclaturally correct names as species of *Triticum*: *T. boeoticum*, *T. monococcum*, *T. thaoudar* and *T. urartu* (Table 1.1, column 1). Note that the name of each species is a binomial; the word(s) in roman type that follows the name of the species refer to the author(s) of the name. In the table, the date when each name was published is shown in parentheses after the author’s name.

There is no nomenclaturally correct name for the “*aegilopoides*” taxon as a distinct species of *Triticum* even though taxonomists now agree that it belongs in *Triticum*. The binomial *Triticum aegilopoides* has been published by two different people, Forsskål (1775) and Körnicke (1885), but they used the name for different taxa. Forsskål applied it to a group of plants that are now

Table 1.1 Names of four close relatives of *Triticum monococcum sensu stricto* under different taxonomic interpretations. The dates show the year that the name was published by the author concerned. Forsskål published the name *Triticum aegilopoides* in 1775, but the specimen that shows what he meant by the name belongs to a different genus

Five equal entities		Two species, <i>T. monococcum</i> with two subspecies (adopted by Slageren [1994])		Three species, <i>T.boeoticum</i> with two subspecies (see Ciaffi <i>et al.</i> 1998)
All species	All subspecies of <i>T.monococcum</i>			
<i>T.monococcum</i> L. (1753)	<i>T.monococcum</i> L. (1753) subsp. <i>monococcum</i>	<i>T.monococcum</i> L. (1753)	<i>T.monococcum</i> L. (1753) subsp. <i>monococcum</i>	<i>T.monococcum</i> L. (1753)
<i>Crithodium aegilopoides</i> Link (1834) (No legitimate name in <i>Triticum</i> at species level for this entity.)	<i>T.monococcum</i> subsp. <i>aegilopoides</i> (Link) Thell. (1918)		<i>T.monococcum</i> subsp. <i>aegilopoides</i> (Link) Thell. (1918)	<i>T.boeoticum</i> Boiss. subsp. <i>boeoticum</i> (effectively 1853) [not <i>T.boeoticum</i> subsp. <i>aegilopoides</i> (Link) E. Schiem. (1939); see text for explanation]
<i>T. boeoticum</i> Boiss. (1854)	<i>T.monococcum</i> subsp. <i>boeoticum</i> (Boiss.) Á. Löve & D. Löve (1961)			
<i>T. thaoudar</i> (Hausskn.) Jakubz. (1932)	<i>T.monococcum</i> subsp. <i>thaoudar</i> (Hausskn.) Flaksb. (Flaksberger 1913)			
<i>T. urartu</i> Thumanjan <i>ex</i> Gandilyan (1972)	<i>T.monococcum</i> subsp. <i>urartu</i> Á. Löve & D. Löve (1961)	<i>T. urartu</i> Thumanjan <i>ex</i> Gandilyan (1972)		<i>T. urartu</i> Thumanjan <i>ex</i> Gandilyan (1972)

known as *Elionurus hirsutus* (Clayton et al. 2002) that do not belong in the *Triticeae*. Körnicke applied it to the same group of plants that Link called *Crithodium aegilopoides* but, according to the Code (McNeill et al. 2006), if there have been two different uses of a name, its meaning is fixed by its first use, no matter how incorrect later taxonomists may find that use to be. This means that Körnicke's use was illegitimate.

If one wants to mention both uses of the name in an article, one can distinguish between them by the author citation: *T. aegilopoides* Forssk. is the taxon now included in *Elionurus*; *T. aegilopoides* (Link) Balansa ex Körn. refers to the taxon that was first described and named by Link; this is indicated by the presence of his name, in parentheses, immediately after the scientific name. Link's name is followed by the abbreviation of the name of the person who published the new combination, Körnicke. In doing so, however, Körnicke acknowledged that he was publishing the combination based on information provided by Balansa; this is indicated by the "Balansa ex" portion of the authorship. *Ex* is Latin for "from within" or "out of" so the citation is saying that Körnicke obtained information from within Balansa's work that made him decide to publish the combination *Triticum aegilopoides*. Balansa may have made notes on herbarium specimens or discussed the matter with Körnicke, but he did not formally publish the name.

In most papers, there is no scientific need to cite the author of the scientific names used although it is usually an editorial requirement. It is, for instance, hard to imagine a paper, other than one about nomenclature, in which both uses of *Triticum aegilopoides* would be employed. If the "*aegilopoides*" taxon is to be named as a species of *Triticum* distinct from the other four species, a different epithet must be used. So far as we know, no one has published such a name.

Each of the five entities discussed in this section has also been named as a subspecies of *Triticum monococcum* (Table 1.1, column 2), but most scientists no longer consider this treatment appropriate. For instance, Slageren (1994) concluded that the "*urartu*" epithet referred to a taxon that should be recognized as a species whereas the "*aegilopoides*", "*boeoticum*", and "*thaoudar*" epithets referred to taxonomically insignificant variants of a single taxon that are best treated as one of two subspecies of *T. monococcum*, the other subspecies being *T. monococcum* subsp. *monococcum*. The correct name at the subspecies level for the combined "*aegilopoides-boeoticum-thaoudar*" entity is *T. monococcum* subsp. *aegilopoides*, because the Code (Art. 11) states that, when combining species or lower ranks, the correct epithet is the one that was first used at the desired rank, in this case subspecies.

This is not the end of the story. Ciaffi et al. (1997, 1998) argued, on the basis of information presented in their papers, that three of the entities should be recognized as species: *T. monococcum*, *T. boeoticum*, and *T. urartu*. They stated that the "*thaoudar*" entity should be treated as a subspecies of *T. boeoticum*, for which they used the name *T. boeoticum* subsp. *thaoudar*. This name was published by Grossheim in 1939 and is nomenclaturally correct, but the name of the other subspecies should be *T. boeoticum* subsp. *boeoticum*, not *T. boeoticum*

subsp. *aegilopoides*, as used by Ciaffi et al. The reason for this is that Ciaffi et al. state: "Two morphological types of *T. boeoticum* are recognized: subsp. *aegilopoides*. . . and subsp. *thaoudar*. . ." (p. 124). Whenever a species is subdivided, one automatically creates a subdivision that includes the type specimen of the species. The name of this subdivision repeats the specific epithet, in this case, "*boeoticum*". (The type specimen is the specimen, usually chosen by the name's author, that anchors the meaning of a name). Thus, the correct name of the second subspecies recognized by Ciaffi et al. is *T. boeoticum* subsp. *boeoticum*. Although "*aegilopoides*" was the first epithet used at the subspecies level, the autonym (i.e., the automatically generated name) has priority. Note that the primary findings of Ciaffi et al. are not affected; one merely needs to substitute *T. boeoticum* subsp. *boeoticum* for *T. boeoticum* subsp. *aegilopoides* when reading their papers.

There is one more wrinkle to this story. The simplest concerns the spelling of "*boeoticum*". Boissier (1854) used two different spellings for this epithet, "*boeoticum*" and "*baeoticum*". Because the type specimen was collected in Boeotia, Greece, the correct spelling is "*boeoticum*".

In the above paragraphs, we showed how different taxonomic treatments resulted in taxa having multiple nomenclaturally correct names. Equally importantly, it means that some names have more than one meaning. For instance, the name *Triticum monococcum* may refer to only one of the five entities treated; to all but the "*urartu*" entity, as in Slageren (1994); or to all five of them (because it was the only one in existence until Link published *Crithodium aegilopoides*). How can one determine which usage an author was adopting? Sometimes it is made clear by the other names used, as in the article by Ciaffi et al. (1998). Another approach, and one that will enable people to interpret what group is intended even if there are further changes in the taxonomic treatment of these entities after publication of a paper, is to prepare herbarium specimens from representative mature plants of the material used in the study and deposit these voucher specimens in a herbarium that will loan them to other researchers on request. A third approach is to cite the taxonomic treatment being followed, but this will not resolve all problems. For instance, if an author follows Slageren, it may be impossible to determine, without looking at voucher specimens, whether references to *T. monococcum* subsp. *aegilopoides* mean *T. boeoticum* subsp. *boeoticum* or *T. boeoticum* subsp. *thaoudar*.

The above paragraphs are about nomenclature. What is not discussed is which of the three taxonomic treatments is best, nor the criteria to be used in deciding what is "best". Those are the questions that must be answered if a single taxonomic treatment is to be adopted. We are recommending acceptance of the treatment based on the findings of Ciaffi et al. (1997, 1998), because their treatment best reflects what is now known about the genetic relationships within the complex. If, however, a paper is published in which van Slageren's treatment is followed, one needs to bear in mind that statements about *T. monococcum* subsp. *aegilopoides* may be referring to either, or both, *T. boeoticum* subsp. *boeoticum* and *T. boeoticum* subsp. *thaoudar*.

1.4 Taxonomic Treatment in this Chapter

As stated earlier, there are no universally accepted rules for making taxonomic decisions. The reproductive behavior of plants tends to make a mockery of attempts to develop such rules. For instance, a frequently heard suggestion is that genera should be monophyletic. This means that a genus should include all species derived from the same common ancestor. This works well in groups where diversification is primarily the result of lineage splitting; it is difficult to follow in a group such as the *Triticeae* that exhibits complex patterns of reticulation and polyploidy.

Another approach to generic classification that has been strongly advocated for the *Triticeae* (Löve 1984; Dewey 1984; Yen et al. 2005a) is to base generic circumscription on genomic composition. Difficulties with strict application of this approach include our ignorance of the genomic composition of several species and the imperfect correlation between morphology and genomic composition. Nevertheless, to the extent that it has been investigated, genomic composition tends to indicate the groups of species to which the primary ancestors of allopolyploid species belong (Svitashev et al. 1996; Mason-Gamer 2001; Liu et al. 2006; Sun et al. 2006). It may, however, underestimate the contribution of hybridization to the tribe's diversity.

Still another approach emphasizes the value of morphologically distinguishable genera. One problem with this approach is that no taxonomist is equally familiar with all the species of the tribe. Groups of species may be evidently distinct to those familiar with them, but indistinguishable to those encountering them for the first time. The ease and low cost of sharing digital images should reduce the extent to which this differing familiarity is a problem in the future. Another question is whether all morphologically distinguishable species groups should be recognized as genera. If they are, *Sitanion* Raf. undoubtedly merits recognition at the generic level.

Taxonomists rely heavily on each other's work when preparing generic descriptions or prepare descriptions based solely on the species occurring in their region of interest. The first approach tends to preclude the adoption of new characters in delimiting genera. It may also lead to the perpetuation of errors. The value of regional descriptions depends on how much of the diversity within the genus is represented in the area concerned. Developing useful circumscriptions of new generic interpretations requires developing detailed species descriptions for a high proportion of the species in each genus.

An additional concern with generic delimitation in an economically important group such as the *Triticeae* is tradition. Changes in generic circumscription should be made if the resulting classification is a substantially closer fit to the evolutionary history of a group. On the other hand, if what is involved is breaking a well-known, easily identified genus into two or more sister genera, or the combination of two long-established sister genera into a single genus, the

gain in understanding conveyed by the closer fit to evolutionary history needs to be weighed against its impact on our ability to interpret existing literature.

In proposing any new generic treatment, consideration should be given to providing others with the information needed to apply it. In general, this requires providing more detailed morphological information than required by the Code (McNeill et al. 2006). Indeed, the Code only requires that one state how a new taxon differs from those with which it might be confused. This means that the Code's requirements for describing a new genus can be met by stating how its genomic constitution or its sequence for a particular gene differs from that found in other genera. In practice, people usually include information about the distinctive morphological attributes of the new genus.

Making decisions at the species level presents different problems, one being that very little is known about most of the species. What is known is that many of the taxa recognized as species will, on occasion, hybridize and form viable seed. If these seeds germinate and yield fertile, or partially fertile, plants, then backcrossing to one or both parents is highly likely. This tends to lead to a blurring of morphological distinctions. If this occurs on a wide scale, the two species are usually reduced in rank or completely disregarded. The problem is deciding at what point hybridization and backcrossing is so common that recognizing two taxa is not justified.

Knowledge of the putative taxa in their native habitats is critical. Taxa that are morphologically very similar may occupy distinct ecological niches, something that is hard to discern from herbarium specimens or plants grown on experimental farms. Growth under experimental conditions may, however, reveal that some differences used to distinguish taxa are determined by environmental conditions. For instance, in the *Bromus carinatus* Hook. and Arn. complex, development of hairy leaf sheaths and culms is controlled by exposure to light (Amme, pers. comm. to Barkworth).

Cultivated species present another suite of taxonomic problems. At least in the past, there was a tendency for those working with cultivated taxa to recognize all morphologically distinct strains as distinct species, one of the arguments being that binomials are easier to remember than the trinomials required if the taxa are recognized as subspecies or varieties. Another tendency is to treat all cultivated taxa with the same genomic constitution as a single species, without recognizing any infraspecific taxa. Alternatively, plants with similar morphological attributes may be placed in the same taxon even though the features may be a consequence of several generations of selection from different seed stocks.

1.4.1 Taxonomic Treatment in this Chapter: The Genera

In our generic decisions (Table 1.2), we have, like Yen et al. (2005a), tried to balance tradition with the desire for a taxonomic treatment that reflects

Table 1.2 Genera of the *Triticeae*, and their level of acceptance. Names for generic characterizations accepted in this chapter are in **boldface**; names of other characterizations are in *italics*. Genome symbols follow the recommendations of the International *Triticeae* Consortium (<http://herbarium.usu.edu/Triticeae/>). Modifications of the basic genome are ignored. If two symbols are separated by a comma (e.g., **St, H, E**), at least two of the genomes are found in all the species; if they are written together (e.g., **StY**) all species contain both genomes, some species may contain more than one copy of either or both genomes. Note: Barley breeders use **H**, not **I**, for the genome in cultivated barley. The generic treatments listed include keys and descriptions

Name	Comments
Level of acceptance	Morphological characterization followed by additional information. (Generic treatment).
Genomic composition	
<i>Aegilops</i>	Annual; inflorescences often with highly reduced, sterile spikelets basally and distally, nodes with 1 spikelet; disarticulation in the rachis, spikelets usually attached to the internode above, sometimes with the internode below, or disarticulation below the basal fertile spikelet; spikelets with 2 or more florets; glumes ovate to rectangular or lanceolate, backs rounded, tip toothed or awned; lemmas not keeled, 1–3-toothed or awned.
Usual use, accepted here.	
B, C, D, M, N, U, X	The B and D genomes of <i>Aegilops</i> are also found in <i>Triticum</i> allopolyploids. Species of <i>Aegilops</i> with these genomes will hybridize with cultivated <i>Triticum</i> species, including <i>T. aestivum</i> , the most widely cultivated species. We follow the usual floristic practice in keeping the two genera as distinct. Yamane and Kawahara (2005) show that <i>Aegilops</i> , <i>Secale</i> , and <i>Triticum</i> formed a single clade but, if <i>A. speltoides</i> were placed in a separate genus (which would be called <i>Sitopsis</i>), each of the other three genera would be monophyletic on their tree (but see also Sallares and Brown 2004; Petersen et al. 2006). Here we retain the traditional interpretation, retaining <i>A. speltoides</i> in <i>Aegilops</i> . (Slageren 1994)
<i>Agropyron</i>	Perennial, usually tufted; inflorescence nodes with 1 spikelet, internodes < 1/3 spikelet length, disarticulation below florets; spikelets with 2 or more florets; glumes keeled, with wide scarious margins, not or shortly awned; lemmas keeled, not or shortly-awned; paleas as long as or slightly shorter than the lemmas; anthers >3 mm.
Universally accepted for “crested wheatgrasses”.	<i>Agropyron</i> is now universally restricted to the “crested wheatgrasses” such as <i>A. cristatum</i> . An interdisciplinary systematic study of the genus is needed to evaluate the numerous specific and infraspecific taxa that have been recognized.
P	
<i>Amblyopyrum</i> (Jaub. Spach) Eig	Plants annual; inflorescences long (>15 cm), nodes with 1 spikelet, disarticulating in the rachis, at base of internode below the spikelet; spikelets appressed, shorter than or slightly longer than the adjacent internode; glumes coriaceous, rounded on the back, unawned; lemmas unawned; paleas slightly shorter than the lemmas; anthers about 4 mm.
Sometimes included in <i>Aegilops</i> .	<i>Amblyopyrum</i> is native to Turkey and Armenia. It is sometimes included in <i>Aegilops</i> , sometimes recognized as a distinct genus. Phylogenetic analyses have been contradictory, some showing it as basal and related to <i>Aegilops speltoides</i> Tausch, others showing it as terminal (Yamane and Kawahara 2005 and references therein). Its genome structure may reflect a history of both hybridization and introgression (Sallares and Brown 2004). (Slageren 1994).
T	

Table 1.2 (continued)

Name	Comments
Level of acceptance	Morphological characterization followed by additional information. (Generic treatment).
Genomic composition	
<i>Anthosachne</i> Steud. Accepted by Yan and Yang (1999), Yen et al. (2005a); not accepted by Edgar and Connor (2000), nor Wheeler et al. (2002).	Perennial; culms loosely clumped; inflorescences with 1 spikelet per nodes; spikelets distant, usually >1.5 times the length of the internodes, disarticulation below the florets; glumes not keeled; lemmas not keeled, usually with awns equal to or longer than the lemma body. <i>Anthosachne</i> used to include both Asian and Australasian taxa but, as currently interpreted, it is restricted to Australia and New Zealand. With this restricted interpretation, Connor and Jacobs support its recognition (pers. comm. to Barkworth, 2008).
StHW	
<i>Australopyrum</i> (Tzvelev) Á. Löve Accepted by Yan and Yang (1999), Yen et al. (2005a), Wheeler et al. (2002) and Edgar and Connor (2000).	Plants perennial, caespitose; inflorescences with 1 spikelet per node; spikelets >3 times as long as the rachis internodes; disarticulation below the florets; glumes keeled; lemmas shortly awned. <i>Australopyrum</i> is native to Australia, New Zealand, and New Guinea. The Australian and New Zealand species are diploids. The W genome is found in all species of <i>Australopyrum</i> and <i>Anthosachne</i> . The New Guinea species is known only from a few herbarium specimens; it has not been examined cytologically.
W	
<i>Crithopsis</i> Accepted	Annuals; inflorescence dense, with 2 spikelets per node, spikelets >4× length of the adjacent internodes, with two florets, the lower floret bisexual, the upper floret sterile and reduced; glumes strap-shaped, coriaceous; lemmas of the bisexual florets shortly awned; anthers 4–7 mm. <i>Crithopsis</i> includes only one species. It grows from the eastern Mediterranean to southwest Asia.
K	
<i>Dasypyrum</i> Accepted	Annual or perennial; inflorescences with 1 spikelet per node; disarticulation in the rachis, below the spikelets; spikelets >2.5 times as long as the rachis internodes, with 2–4 florets; glumes and lemmas shortly awned; anthers 5–7 mm. <i>Dasypyrum</i> includes two species. It grows in the Mediterranean region and southeastern Europe. (Frederiksen 1991a).
V	
<i>Douglasdeweya</i> Yen et al. (2005b).	<i>Douglasdeweya</i> differs from <i>Pseudoroegneria</i> , donor of the St genome, in having strongly keeled glumes and lemmas, and tapered glumes with narrow translucent margins, and from <i>Agropyron</i> , donor of the P genome, in having spikelets that are <1 to slightly more than twice as long as the rachis internodes. Both known species are rhizomatous.
StP	
<i>Elymus</i> (widest) Rarely accepted	<i>Douglasdeweya</i> is known primarily from plants grown from two genebank accessions that have been grown on experimental plots. Its range includes northwestern Iran and the Karachay-Cherkess Republic, Russia. In describing the genus, Yen et al. (2005b) did not state how it differs from <i>Elymus s.str.</i>
Two or more of E, H, Ns, P, St, W, Y, and L	Perennial or annual, tufted or strongly rhizomatous; inflorescence disarticulating or not, occasionally paniculate; nodes with 1–many spikelets, spikelets more than twice the length of the adjacent internode; glumes lanceolate-ovate to awnlike, sometimes absent or highly reduced, sometimes keeled; lemmas awned or unawned; paleas distinctly shorter to slightly longer than the lemmas, ovate to lanceolate; anthers 1–9 mm.

Table 1.2 (continued)

Name	Comments
Level of acceptance	Morphological characterization followed by additional
Genomic composition	information. (Generic treatment).
<p><i>Elymus</i></p> <p>As adopted here</p> <p>St plus at least one of H, Y, W or an unknown genome.</p>	<p>A few contemporary taxonomists treat <i>Elymus</i> as including <i>Taeniatherum</i> and almost all the perennial genera other than <i>Agropyron</i> and <i>Hordeum</i>.</p> <p>Perennial, tufted or strongly rhizomatous; inflorescence disarticulating or not; nodes with 1–5 spikelets, internodes <1/2 as long the spikelets; glumes lanceolate to lanceolate-ovate or awnlike, sometimes absent or highly reduced, if lanceolate to lanceolate-ovate, usually flexible and weakly keeled, sometimes strongly keeled distally; lemmas awned or unawned; paleas subequal to the lemmas, lanceolate to lanceolate-ovate; anthers 1–5 mm.</p>
<p><i>Elymus s.str.</i></p> <p>The narrowest interpretation, adopted by Yen et al. (2005a).</p> <p>StH, StStH, StHH</p>	<p>This is our most controversial decision. It reflects the importance we attach to there being current floristic treatment of a group. For this reason, we include <i>Roegneria</i>, but not <i>Kengyilia</i>, in <i>Elymus</i>. We also include the Australasian hexaploids and octoploids, pending a revision of these species that includes descriptions. There are some allopolyploid South American species currently included in <i>Elymus</i> that are known to possess the H genome and one other genome that is not the St genome (Linde-Laursen and Seberg 2001). Until more information is available, both morphological and genomic, they are retained in <i>Elymus</i>.</p> <p>Perennials; tufted or rhizomatous; inflorescence nodes with 1–5 spikelets, internodes <1/2 as long the spikelets; disarticulation below the florets; glumes lanceolate-ovate to lanceolate-ovate or awnlike, sometimes absent or highly reduced, if lanceolate to lanceolate-ovate, usually flexible and weakly keeled, sometimes strongly keeled distally; lemmas awned or unawned; paleas subequal to or slightly longer than the lemmas, lanceolate, keels finely and densely toothed distally, veins converging uniformly to the tip, intercostal membrane usually shorter than to equaling the vein tips; anthers 1–5 mm long.</p> <p>This interpretation differs from the one adopted here primarily in the exclusion of species with the Y and W genomes. Even with this narrow interpretation, <i>Elymus</i> remains morphologically variable; just how variable cannot be determined in the absence of comparable descriptions of all the relevant species.</p>
<p><i>Elytrigia</i></p> <p>Accepted by Tsvelev (1976) and Wu et al. (2006); not by Barkworth et al. (2007).</p> <p>St, E, H, N</p>	<p>Plants perennial, often rhizomatous; inflorescences with 1 spikelet per node; disarticulation below the glumes or below the florets; glumes rectangular to lanceolate, often stiff, often strongly keeled distally; anthers longer than 4 mm.</p> <p>There is little consistency in the interpretation of <i>Elytrigia</i>. It consists of those species that do not fit easily into other genera recognized in a given work. The type species is <i>Elytrigia repens</i>. When, as here, <i>Et. repens</i> is included in <i>Elymus</i>, <i>Elytrigia</i> cannot be used as a generic name.</p>
<p><i>Eremium</i></p> <p>Not accepted in Soreng et al. (2003)</p> <p>N</p>	<p>Perennials; inflorescences with 1–2 spikelets per node; disarticulation below the florets; glumes linear-subulate, longer than the adjacent florets, margins and occasionally the abaxial surface densely long-ciliate, unawned; lemmas shortly awned; anthers 2.2–4.4 mm long.</p>

Table 1.2 (continued)

Name	Comments
Level of acceptance	Morphological characterization followed by additional
Genomic composition	information. (Generic treatment).
<i>Eremopyrum</i> Universally accepted FXe	<p><i>Eremium</i> is a South American genus with one species. It differs from other genera of <i>Triticeae</i> in the ciliate margins of its glumes. Dubcovsky et al. (1997) demonstrated that, genomically, it resembles other species of <i>Leymus</i>, some of whose members it resembles morphologically. (Seberg et al. 1991).</p> <p>Annuals; inflorescence with 1 spikelet per node; disarticulation below the florets; spikelets $\geq 3 \times$ the length of the internodes; glumes coriaceous, keeled; lemmas coriaceous, keeled distally; anthers 0.4–1.3 mm.</p> <p><i>Eremopyrum</i> is a genus of 5–10 species that grow in steppes and semi-desert regions from Turkey to Central Asia and Pakistan. (Frederiksen 1991b).</p>
<i>Festucopsis</i> Accepted L	<p>Perennials; sheaths disintegrating into more or less parallel fibers; inflorescence with 1 spikelet per node; disarticulation below the florets; glumes coriaceous with a narrow membranous margin, unawned; lemma coriaceous, upper lemmas usually awned; anthers 3–3.5 mm.</p> <p><i>Festucopsis</i> is endemic to Albania where it grows on serpentine soils. Its genome may be present in some species of <i>Psammopyrum</i> that are included here in <i>Thinopyrum</i> (Ellneskog-Staam et al. 2003). Tutin et al. (1980) included <i>Peridictyon</i> in <i>Festucopsis</i>, which was first described in 1991. (Seberg et al. 1991).</p>
<i>Henrardia</i> Accepted O	<p>Annual; inflorescences with 1 spikelet node; disarticulation in the rachis; spikelets falling with the internode to the spikelet above; glumes lying side by side, coriaceous, unawned, lying side by side; lemmas membranous, unawned.</p> <p><i>Henrardia</i> has one species. It is native to western Asia.</p>
<i>Heterantherium</i> Accepted Q	<p>Annual; inflorescences with 1 spikelet per node; disarticulation in the rachis, the disarticulating units bearing at least one spikelet with bisexual florets at the base and 1 or more sterile distal spikelets; glumes subulate; lemma of bisexual florets oval, inflated, long-awned; palea keels prolonged, forming awnlike extensions.</p> <p><i>Heterantherium</i> includes only one species. It grows in western and central Asia.</p>
<i>Hordelymus</i> Widely accepted, but possibly better included in <i>Leymus</i> N	<p>Perennial; inflorescences with 3 spikelets per node; disarticulation below the florets; glumes linear-lanceolate, connate below, awned; lemmas long-awned.</p> <p><i>Hordelymus</i> consists of a single species that grows from northern Europe to the Mediterranean and Crimea. Ellneskog-Staam et al. (2006) showed that it contains only the N genome, which suggests that consideration should be given to transferring it to <i>Leymus</i>. It resembles some of the recently transferred species of <i>Hystrix</i> in growing in shady sites.</p>
<i>Hordeum</i> Universally accepted, H, Xa, Xu, I	<p>Annual or perennial; inflorescences disarticulate (except cultivated derivatives); nodes with 3 spikelets, lateral spikelets pedicellate; internodes often much shorter than spikelets; spikelets with 1 floret; glumes awnlike; florets of lateral spikelets reduced and sterile (except cultivated derivatives); lemmas of fertile florets long-awned.</p>

Table 1.2 (continued)

Name	Comments
Level of acceptance	Morphological characterization followed by additional
Genomic composition	information. (Generic treatment).
<p><i>Hystrix</i> Accepted by Wu et al. (2006) and Bothmer et al. (2005); not by Barkworth et al. (2007). StH or N</p>	<p>The H genome of <i>H. jubatum</i> and some other species is also found in many allopolyploids. The I genome is confined to <i>H. vulgare</i> and <i>H. bulbosum</i>, the Xa and Xb genomes to <i>H. murinum s.l.</i> and <i>H. marinum s.l.</i>, respectively. <i>Critesion</i> was at one time used for species with the H, Xa, and Xu genomes (Löve 1984; Barkworth and Dewey 1985), but now all scientists include them in <i>Hordeum</i>. (Bothmer et al. 1995)</p> <p>Perennial, tufted; inflorescence nodes with 2–many spikelets; disarticulation below the florets; glumes absent or highly reduced; lemmas awned or unawned; paleas subequal to the lemmas, lanceolate.</p> <p><i>Hystrix</i> differs from other genera in its lack of glumes, but there are species of <i>Leymus</i> that have never been included in <i>Hystrix</i>, e.g., <i>L. salinus</i>, that also have highly reduced glumes. An additional problem is that some of its members, including the type species, are StH tetraploids, whereas others have two versions of the N genome. A nomenclatural problem is that if the type species is included in <i>Elymus</i>, as here, <i>Hystrix</i> cannot be used as a generic name. This creates a problem because some of the species that have been placed in <i>Hystrix</i> do not yet have a name in any other genus. (Baden et al. 1997)</p>
<p><i>Kengyilia</i> Accepted by Wu et al. (2006). StPY</p>	<p>Perennial, sometimes rhizomatous; inflorescences usually with 1 spikelet per node, sometimes with 2; spikelets usually closely spaced; disarticulation below the florets; glumes not keeled or only keeled distally; lemmas rarely keeled, densely pilose or hirsute, unawned or shortly awned; paleas retuse, obtuse, or 2-lobed.</p> <p>Species of <i>Kengyilia</i> grow at high elevations in central Asia and on the Tibetan plateau. It is accepted here in part because of its acceptance in the <i>Flora of China</i>.</p>
<p><i>Leymus</i> Almost universally accepted, not always as broadly as here. N</p>	<p>Perennials, often rhizomatous; inflorescences usually spikelike, paniculate in some species, if spikelike with 1–several spikelets per node; disarticulation below the florets; glumes absent or broadly to very narrowly lanceolate, flexible and not keeled if broadly lanceolate, stiff and keeled if narrowly lanceolate, unawned; lemmas not or shortly awned; anthers usually >3 mm.</p> <p>Most species of <i>Leymus</i> grow in open habitats with alkaline soils, but some species that lack glumes grow in forested, neutral to acidic soils. The species that lack glumes are sometimes included in <i>Hystrix</i>. Wang and Jensen (1994) demonstrated that <i>Leymus</i> did not, as postulated by Löve (1984) contain the J genome.</p>
<p><i>Pascopyrum</i> Accepted in Barkworth et al. (2007) St, H, N</p>	<p>Perennial, rhizomatous; inflorescences with 1 spikelet at most or all nodes; spikelets at more than 2× the length of the adjacent rachis internode; disarticulation below the florets; glumes tapering from midlength or below, keeled; lemmas shortly awned; anthers 2.5–6 mm.</p> <p><i>Pascopyrum</i> is an octoploid that combines the genomes of <i>Elymus s.str.</i> with those of <i>Leymus</i>. It is restricted to North America. It has only one species.</p>

Table 1.2 (continued)

Name	Comments
Level of acceptance	Morphological characterization followed by additional information. (Generic treatment).
Genomic composition	
<i>Peridictyon</i>	Perennial; leaf sheaths disintegrating into a reticulate fibers; inflorescences with 1 spikelet per node; disarticulation below the florets; glumes lanceolate, usually shortly awned; lemmas coriaceous, shortly awned; anthers >4 mm.
Accepted	<i>Peridictyon</i> has one species that grows on calcareous rocks in Bulgaria and Greece. It was described in 1991, after publication of <i>Flora Europaea</i> (Tutin et al. 1980) where it was included in <i>Festucopsis</i> .
Xp	
<i>Psathyrostachys</i>	Perennial; inflorescences with 2–3 spikelets per node; disarticulation in the rachis nodes above the spikelet, this sometimes tardy; spikelets 2–3 times the length of the internodes; glumes short, subulate, coriaceous; lemmas acute to awned; anthers 3.3–7 mm.
Accepted	<i>Psathyrostachys</i> has eight species all of which are native to arid regions of central Asia. Bödvarsdóttir and Anamthawat-Jónsson (2003) found no molecular probes that would distinguish the chromosomes of <i>Psathyrostachys</i> from those of <i>Leymus</i> . The two genera are morphologically similar. (Baden 1997).
N	
<i>Pseudoroegneria</i>	Perennial, usually tufted, sometimes shortly rhizomatous; inflorescence nodes with 1 spikelet, internodes about ½ spikelet length, disarticulation below florets; glumes lanceolate, usually not keeled, essentially unawned; lemmas awned or unawned, awn strongly curved; paleas equaling the lemmas, narrowing distally; anthers >3 mm.
Accepted by Wu et al. (2006) and Barkworth et al. (2007); not accepted by Tutin et al. (1980), nor by Tsvelev (1976).	<i>Pseudoroegneria</i> grows in Asia and North America. Its genome is the most pervasive in the tribe, being found in almost all species of <i>Elymus</i> (including <i>Roegneria</i>), <i>Anthosachne</i> , and <i>Kengyilia</i> and some species of <i>Thinopyrum</i> .
St	
<i>Roegneria</i>	Plants perennial, not rhizomatous; inflorescences with 1 spikelet per node; disarticulation below the florets; glumes lanceolate to lanceolate-ovate; lemmas awned or unawned; paleas often conspicuously shorter than the lemmas, often (?always) widest beyond midlength, keels distally scabrous, the teeth more widely separated than in StH species.
Accepted by Yen et al. (2005a).	Yen et al.'s (2005a) interpretation of <i>Roegneria</i> is considerably narrower than that of Baum et al. (1991), being explicitly restricted to species having only the St and Y genomes. Neither interpretation was accepted in Wu et al. (2006), although most of its species are native to China. For that reason, it is not recognized in this volume. There is a lack of the morphological information needed to distinguish <i>Roegneria</i> from <i>Elymus</i> . Salomon and Lu (1992) suggested that the palea characters mentioned might serve this purpose, but they need to be examined in more species.
St, Y	
<i>Secale</i>	Annual or perennial; inflorescences with 1 spikelet per node, disarticulation beneath the florets or in the rachis; glumes strongly keeled, awned; lemmas keeled, keels ciliate, terminating in a strongly scabrous awn; anthers 2.3–12 mm.
Universally accepted	
R	

Table 1.2 (continued)

Name	Comments
Level of acceptance	Morphological characterization followed by additional information. (Generic treatment).
Genomic composition	
<i>Sitanion</i>	<i>Secale</i> is native to the Mediterranean region and western Asia. (Frederiksen and Peterson 1998).
No longer accepted	Perennials; inflorescences with 2 or more spikelets per node; disarticulation initially in the rachis, spikelets falling with the rachis internode below; spikelets $\geq 2.5\times$ the length of the length of the rachis internodes; glumes subulate or divided into subulate divisions; lemmas long-awned; anthers 1–5 mm.
StH	<i>Sitanion</i> includes 2 western North American species that are strongly adapted for wind dispersal. They are distinctive, which is why they were recognized as a separate genus. They are now included in <i>Elymus</i> , because they are genetically closely related to other species of that genus. Hybrids with other species of <i>Elymus</i> are common; they are usually sterile. (Wilson 1963).
<i>Stenostachys</i>	Perennial, stoloniferous; culms loosely clumped; inflorescences with 1 spikelet per node; spikelets about 1.5 times the length of the internodes; disarticulation below the florets; glumes absent or narrow or subulate, comparable in length to the lowest lemma; lemmas weakly to strongly keeled, awned, awn shorter than the lemma body.
Accepted by Edgar and Connor (2000).	<i>Stenostachys</i> is endemic to New Zealand. Stewart et al. (2005) demonstrated that its members are HW tetraploids. (Connor 1994)
HW	
<i>Taeniatherum</i>	Annuals; inflorescences usually with 2 spikelets; disarticulation below the florets; glumes subulate, connate basally; lemmas long awned, becoming divergent at maturity; anthers 0.8–1 mm.
Universally accepted	<i>Taeniatherum</i> is native around the Mediterranean and in western Eurasia. (Frederiksen 1986).
Ta	
<i>Thinopyrum</i>	Perennial; inflorescences with 1 spikelet per node; spikelets distant; disarticulation at the base of the spikelets; glumes stiff, not keeled or keeled distally, acute to truncate or retuse; lemmas usually unawned; anthers >3 mm long.
Including	<i>Thinopyrum</i> is a poorly known genus. All its members have the E genome or a modified version that is sometimes identified as J ; allopolyploid species have the St genome plus, in some instances, either the L genome of <i>Festucopsis</i> or the P genome of <i>Agropyron</i> (Ellneskog-Staam et al. 2003). This relatively simple picture was complicated by the discovery that, in some species, the E and P genome chromosomes have centromeric regions that hybridize with St DNA probes (Chen et al. 1998; Refoufi et al. 2001).
<i>Psammopyrum</i>	
Accepted in Barkworth et al. (2007); not accepted in Tsvelev (1976), Tutin et al. (1980), and Wu et al. (2006)	
E sometimes with P , St , or L	
<i>Triticum s.l.</i>	Annual; inflorescences disarticulate (except cultivated derivatives); nodes with 1 spikelet; spikelets with 2 or more florets, distal and basal spikelets sometimes greatly reduced and sterile; glumes ovate to rectangular or lanceolate; lemmas awned or long-awned.
Rarely accepted	
(includes <i>Aegilops</i> and <i>Amblyopyrum</i>)	
A, B, C, D, M, N, T, U, X	There are arguments for and against interpreting <i>Triticum</i> in this broad sense, but a strong tradition for not doing so. We recommend the traditional practice.

Table 1.2 (continued)

Name	Comments
Level of acceptance	Morphological characterization followed by additional information. (Generic treatment).
Genomic composition	
<i>Triticum</i>	Annual; inflorescences disarticulation in the rachis (except cultivated derivatives), spikelets usually falling with internode below; nodes with 1 spikelet; spikelets with 2 or more florets, all spikelets similar; glumes ovate to rectangular or lanceolate, keeled; lemmas keeled, short- to long-awned.
Usual use, accepted here.	The A genome is present in all species traditionally placed in <i>Triticum</i> and none of those traditionally placed in <i>Aegilops</i> . The B and D genomes present in polyploid wheats are derived from species of <i>Aegilops</i> .
A, AB, AAB, ABD	

evolutionary history. We have also been influenced by the existence, or lack thereof, of viable keys to the genera and species. For this reason, we differ from Yen et al. in our tendency to favor taxonomic treatments adopted in recent floras that cover large areas. At their best, these provide morphological keys, descriptions, and illustrations (e.g., Wu et al. 2006; Wu and Raven 2007; Barkworth et al. 2007); at a minimum, they provide identification keys and some additional notes (Tsvelev 1976). One advantage of this floristic bias is that, for the vast majority of species, the treatment agrees with current floristic treatments and nomenclaturally correct names are available for almost all the species. In our generic summary (Table 1.2), we indicate the reason for discrepancies between our treatment and those of other authors.

The generic descriptions are drawn from our own knowledge plus information in Tutin et al. (1980), Cope (1982), Davis (1985), Edgar and Connor (2000), Wu et al. (2006), Barkworth et al. (2007). We have made little attempt to create parallel descriptions; instead, we have tried to include the distinctive features of each genus in its description.

1.4.2 Taxonomic Treatment in this Chapter: The Species

Making taxonomic decisions at the species level involves a different set of problems from those involved in making decisions at the generic level. One of these is the vastly different levels of information that exist for different species. For some species only a single, often old, herbarium specimen is available; for others there is information from growth under different environmental conditions, interpopulation and interspecific hybridization, developmental studies, and DNA studies of many portions of their genomes.

Another problem with species delimitation in the *Triticeae*, particularly the perennial species, is the tendency for hybrids to exhibit at least a low level of fertility. The establishment of such hybrids, and their subsequent

backcrossing to one or both parents, leads to confusing patterns of morphological variation. This happens even among species that are described as self-pollinating; almost all such species are, in fact, only predominantly self-pollinating. Cross-pollination, although infrequent, may occur. This can lead to the formation of local populations with a distinctive morphology and ecology. Opinions vary as to whether such entities should be described as distinct species. The problem is exacerbated by the fact that hybrid populations with the same parents may vary from being almost fully fertile to being almost infertile. The treatment of wild species generally follows that of a recent major floristic treatment for their native range; the treatment of cultivated entities reflects an attempt to balance the goal of summarizing what is known about these entities with that of having somewhat equivalent treatment of the cultivated and wild entities.

1.5 Nomenclatural Web Sites

There is no Web site that lists all names that have been used for members of the *Triticeae* together with all the alternative interpretations of each name. The best resource for such information is the database developed for and by curators at the Missouri Botanical Garden, <http://mobot.mobot.org/W3T/Search/vast.html>. This provides the most complete unbiased listing of vascular plant names and synonyms currently available. Because it is designed primarily for the garden's curators, its coverage is most complete for those areas in which the garden has active research programs such as China, the Americas, and parts of southern Asia. Another feature of the site is that many of the names are linked to an image of their publication information.

Another Web site for nomenclatural information is <http://www.ipni.org>, which presents information in the International Plant Names Index (IPNI). It is being developed through collaboration between the Royal Botanic Gardens, Harvard University, and the Australian National Herbarium. It provides information on synonyms that are based on the same type specimen, but not those that reflect a taxonomic judgment. One difficulty in using it is that, because it brings together information from three, independently developed, print catalogs, there are many instances where there is more than one citation given and it is not clear which is correct. Those responsible for the site are working to eliminate such confusing entries, but the process requires checking all the relevant literature, a task that is time-consuming.

There are also, as the curators of Tropicos and the IPNI readily admit, errors present on both sites. These are corrected when they are brought to the curators' attention, but this is not always a simple task. During the preparation of this paper, a question arose concerning the information given for *Triticum thaoudar*. Answering the question involved two separate institutions and resulted in a correction to some of the entries on the two sites.

Nevertheless, the existence of the two sites means that scientists working today have easy, rapid access to much more, and more accurate, nomenclatural information than any of their predecessors, even predecessors working in major taxonomic institutions.

The Royal Botanic Gardens, Kew, has a Web site that lists the names of all genera and species of grasses, and their synonyms, *as recognized by the taxonomists at Kew*. The names are linked to descriptions that are drawn from a database that reflects their taxonomic treatment although much of its information is derived from regional floras that may have employed a different treatment (see Simon 2007).

The International *Triticeae* Consortium is starting to create a Web site that will include some nomenclatural information (see <http://utc.usu.edu/Triticeae>); it will try to supplement, not duplicate, the information provided by Tropicos and IPNI. Development of the Consortium's Web site is, however, dependent on funding.

Acknowledgments It is, fortunately, rarely evident how much work is involved in addressing nomenclatural problems, particularly those associated with older names. For this paper, we chose what seemed like simple examples of the nomenclatural consequences of taxonomic changes. They turned out to be more complex than we thought. Resolution of the problems lead to a flurry of exchanges and a few changes in the nomenclatural Web sites cited. For instance, it was discovered that the combination *Triticum thaoudar* was first made by Nevski, not Boissier, that Haussknecht (1899) based recognition of the “*thaoudar*” entity on his own observation, not Reuter's and that he treated it as a race, a rank that is no longer recognized in the Code (McNeill et al. 2006). None of these discoveries will have a significant impact on current research findings; they do affect the author citation for the names associated with the “*thaoudar*” entity.

We thank Dr. Kathleen Capels for her careful review of an earlier edition of this manuscript, in the course of which she raised nomenclatural questions that could not be answered with the resources available at Utah State University; Drs. Gerrit Davidse, Anna Filatenko, Kanchi Gandhi, Helmut Knüpfner, Robert Soreng, and John Wiersema, all of whom helped resolve the nomenclatural questions raised by Dr. Capels and Drs. Kanchi, Davidse, and Wiersema for amending the IPNI, TROPICOS, and GRIN Web sites to reflect the necessary changes.

Appendix

1. *Aegilops* L.
2. *Aegilops speltoides* Tausch
3. *Agropyron* Gaertn.
4. *Agropyron cristatum* (L.) P. Beauv.
5. *Agropyron dasystachyum* (Hook.) Scribn.
6. *Agropyron inerme* (Scribn. & J.G. Sm.) Rydb.
7. *Agropyron lanceolatum* Scribn. & J.G. Sm.
8. *Agropyron spicatum* (Pursh) Scribn. & J.G. Sm.
9. *Agropyron subvillosus* (Hook.) E.E. Nelson

10. *Amblyopyrum* (Jaub. & Spach) Eig
11. *Anthosachne* Steud.
12. *Australopyrum* (Tzvelev) A. Löve
13. *Bromus carinatus* Hook. & Arn.
14. *Critesion* Raf.
15. *Crithodium aegilopoides* Link
16. *Crithopsis* Jaub. & Spach
17. *Dasypyrum* (Coss. & Durieu) T. Durand
18. *Douglasdeweya* C. Yen, J.L. Yang, & B.R. Baum
19. *Elionurus* Humb. & Bonpl. ex Willd.
20. *Elymus* L.
21. *Elymus dasystachys* Trin.
22. *Elymus hystrix* L.
23. *Elymus lanceolatus* (Scribn. & J.G. Sm.) Gould
24. *Elymus spicatus* (Pursh) Gould
25. *Elytrigia* Desv.
26. *Elytrigia repens* Desv.
27. *Elytrigia spicata* (Pursh) D.R. Dewey
28. *Eremium* Seberg & Linde-Laursen
29. *Eremopyrum* (Ledeb.) Jaub. & Spach
30. *Festuca* L.
31. *Festuca spicata* Pursh
32. *Festucopsis* (C.E. Hubb.) Melderis
33. *Henrardia* C.E. Hubb.
34. *Heteranthelium* Hochst.
35. *Hordelymus* (Jess.) Harz
36. *Hordeum* L.
37. *Hordeum vulgare* L.
38. *Hystrix* Moench
39. *Hystrix californica* (Bol. ex Thurb.) Kuntze
40. *Hystrix patula* Moench
41. *Kengyilia* C. Yen & J.L. Yang
42. *Leymus* Hochst.
43. *Leymus salina* (M.E. Jones) Á. Löve
44. *Leymus secalinus* (Georgi) Tzvelev
45. *Pascopyrum* Á. Löve
46. *Peridictyon* Seberg, Fred. & Baden
47. *Psammopyrum* Á. Löve
48. *Psathyrostachys* Nevski
49. *Pseudoroegneria* (Nevski) Á. Löve
50. *Pseudoroegneria spicata* (Pursh) Á. Löve
51. *Roegneria* K. Koch
52. *Secale cereale* L.
53. *Secale* L.
54. *Sitanion* Raf.

55. *Sitopsis* (Jaub. & Spach) Löve
56. *Stenostachys* Turcz.
57. *Taeniatherum* Nevski
58. *Thinopyrum* Á. Löve
59. *Triticeae* Dumort.
60. \times *Triticosecale* Wittm. ex A. Camus
61. *Triticum* L.
62. *Triticum aegilopoides* Balansa ex Forssk.
63. *Triticum aestivum* L.
64. *Triticum boeoticum* Boiss.
65. *Triticum boeoticum* Boiss. subsp. *boeoticum*
66. *Triticum boeoticum* subsp. *aegilopoides* (Link) E. Schiem
67. *Triticum boeoticum* subsp. *thaoudar* (Hausskn.) E. Schiem
68. *Triticum durum* Desf.
69. *Triticum monococcum* L.
70. *Triticum monococcum* L. subsp. *monococcum*
71. *Triticum monococcum* subsp. *aegilopoides* (Link) Thell.
72. *Triticum thaoudar* (Hausskn.) Reut. ex Nevski
73. *Triticum turgidum* ssp. *durum* (Desf.) Husn.
74. *Triticum urartu* Thumanjan ex Gandilyan

References

- Anamthawat-Jonsson, K., Bodvarsdottir, S.K., Bragason, B.T., Gudmundsson, J. and Koebner, R.M.D. (1997) Wide hybridization between wheat (*Triticum* L.) and Lymegrass (*Leymus* Hochst.). *Euphytica* 93, 293–300.
- Avdulov, N.P. (1931) Karyo-systematische Untersuchungen der Familie Gramineae. *Bull. Appl. Bot. Suppl.* 44, 1–428.
- Baden, C. (1997) A taxonomic revision of *Psathyrostachys* (Poaceae). *Nordic J. Bot.* 11, 3–26.
- Baden, C., Frederiksen, S. and Seberg, O. (1997) A taxonomic revision of the genus *Hystrix* (*Triticeae*, Poaceae). *Nordic J. Bot.* 17, 449–467.
- Barkworth, M.E. (2000) Changing perceptions in the *Triticeae*. In: S.W.L. Jacobs and J. Everett (Eds.), *Grasses: Systematics and Evolution*. CSIRO Publishing, Collingwood, Victoria, Australia, pp. 1101–1120.
- Barkworth, M.E. (2006) A new hybrid genus and 12 new combinations in North American grasses. *Sida* 22, 495–501.
- Barkworth, M.E. and Bothmer, R. von (2005) Twenty-one years later: Löve and Dewey's genomic classification proposal. *Czech J. Genet. Plant Breed.* 41 (Special Issue), 3–9.
- Barkworth, M.E., Capels, K.M., Long, S., Anderton, L.K. and Piep, M.B. (Eds.) (2007) *Magnoliophyta: Commelidinae* (in part): *Poaceae* (part 1). *Flora of North America*, vol. 24. Oxford University Press, New York.
- Barkworth, M.E. and Dewey, D.R. (1985) Genomically based genera in the perennial *Triticeae* of North America: Identification and membership. *Amer. J. Bot.* 72, 767–776.
- Barkworth, M.E., McGrew, J. and Zhu, G.-H. (2005) The *Triticeae* in China. *Czech J. Genet. Plant Breed.* 41 (Special Issue), 65–69.
- Baum, B.R., Yen, C. and Yang, J.-L. (1991) *Roegneria*: Its generic limits and justification for its recognition. *Can. J. Bot.* 69, 282–294.
- Bentham, G. (1882) Notes on *Gramineae*. *J. Linn. Soc., Bot.* 19, 14–134.

- Bödvarsdóttir, S.K. and Anamthawat-Jónsson, K. (2003) Isolation, characterization, and analysis of *Leymus*-specific DNA sequences. *Genome* 46, 673–682
- Boissier, P.E. (1854) *Diagnoses Plantarum Orientalium Novarum*, ser. 1, pt. 13. [privately published], Necomi [Como], Italy.
- Bothmer, R. von, Jacobsen, N., Baden, C., Jørgensen, R.B. and Linde-Laursen, I. (1995) *An Ecogeographical Study of the Genus Hordeum*, 2nd ed. Systematic and Ecogeographic Studies on Crop Gene pools No. 7. International Board for Plant Genetic Resources, Rome, Italy.
- Bothmer, R. von, Salomon, B., Enomoto, T. and Watanabe, O. (2005) Distribution, habitat and status for perennial *Triticeae* species in Japan. *Bot. Jahrb. Syst. Pflanzengesch. Pflanzengeogr.* 126, 317–346.
- Brummitt, R.K. and Powell, C.E. (1992) Authors of Plant Names: A List of Authors of Scientific Names of Plants, with Recommended Standard Forms of Their Names, Including Abbreviations. Royal Botanic Gardens, Kew, England.
- Carlson, J.R. (2007) *Pseudoroegneria*. In: M.E. Barkworth, K.M. Capels, S. Long, L.K. Anderton and M.B. Piep (Eds.), *Magnoliophyta: Commelidinae* (in part): Poaceae (part 1). *Flora of North America*, vol. 24. Oxford University Press, New York, pp. 279–282.
- Chen, Q., Conner, R.L., Laroche, A. and Thomas, J.B. (1998) Genome analysis of *Thinopyrum intermedium* and *Thinopyrum ponticum* using genomic *in situ* hybridization. *Genome* 41, 580–586.
- Church, G.L. (1967) Taxonomic and genetic relationships of eastern North American species of *Elymus* with setaceous glumes. *Rhodora* 69, 121–162.
- Ciafffi, M., Dominici, L. and Lafiandra, D. (1997) Gliadin polymorphism in wild and cultivated einkorn wheats. *Theor. Appl. Genet.* 94, 68–74.
- Ciafffi, M., Dominici, L. and Lafiandra, D. (1998) High molecular weight glutenin subunit variation in wild and cultivated einkorn wheats (*Triticeae* spp., *Poaceae*). *Plant Syst. Evol.* 209, 123–127.
- Clayton, W.D., Harman, K.T. and Williamson, H. (2002). *World Grass Species: Synonymy*. <http://www.kew.org/data/grasses-syn.html>. [accessed October, 2007]
- Cope, T.A. (1982) *Flora of Pakistan*, vol. 143. In: E. Nasir and S.A. Ali (Eds), *Flora of Pakistan*. Pakistan Agricultural Research Council and University of Karachi, Pakistan.
- Connor, H.E. (1994) Indigenous New Zealand *Triticeae*, Gramineae. *New Zeal. J. Bot.* 32, 125–154.
- Daubenmire, R.F. (1939) The taxonomy and ecology of *Agropyron spicatum* and *A. inerme*. *Bull. Torrey Bot. Club* 66, 327–329.
- Daubenmire, R.F. (1960) An experimental study of variation in the *Agropyron spicatum* – *A. inerme* complex. *Bot. Gaz.* 122, 104–108.
- Davis, P.B. (Ed.) (1985) *Flora of Turkey and the East Aegean Islands*. vol. 9. University Press, Edinburgh.
- Dewey, D.R. (1982) Genomic and phylogenetic relationships among North American perennial *Triticeae*. In: J.R. Estes, R.J. Tyrl and J.N. Brunken (Eds.), *Grasses and Grasslands*. University of Oklahoma Press, Norman, pp. 51–58.
- Dewey, D.R. (1983) New nomenclatural combinations in the North American perennial *Triticeae* (Gramineae). *Brittonia* 35, 30–33.
- Dewey, D.R. (1984) The genomic system of classification as a guide to intergeneric hybridization in the perennial *Triticeae*. In: J.P. Gustafson (Ed.), *Gene Manipulation in Plant Improvement*. Plenum Press, New York, pp. 209–279.
- Dhaliwal, H.S., Friebe, B., Gill, K.S. and Gill, B.S. (1990) Cytogenetic identification of *Aegilops squarrosa* chromosome additions in *durum* wheat. *Theor. Appl. Genet.* 79, 769–774.
- Dorofeev, V.F., Filatenko, A.A., Migushova, E.F., Udaczin, R.A. and Jakubziner, M.M. (1979) Wheat. vol. 1. In: V.F. Dorofeev and O.N. Korovina (Eds.), *Flora of Cultivated Plants*. Leningrad (St. Petersburg), Russia. Kolos (in Russian). 346 pp.

- Dubcovsky, J., Schlatter, A.R. and Echaide, M. (1997) Genome analysis of Southern American *Elymus* (*Triticeae*) and *Leymus* (*Triticeae*). *Genome* 40, 505–520.
- Edgar, E. and Connor, H.E. (2000) *Flora of New Zealand*, vol. 5. Manaaki Whenua Press, Lincoln, New Zealand.
- Ellneskog-Staam, P., Salomon, B., Bothmer, R. von and Anamthawat-Jónsson, K. (2003) The genome composition of hexaploid *Psammopyrum athericum* and octoploid *Psammopyrum pungens* (Poaceae: *Triticeae*). *Genome* 46, 164–169.
- Ellneskog-Staam, P., Taketa, S., Salomon, B., Anamthawat-Jónsson, K. and Bothmer, R. von (2006) Identifying the genome of wood barley *Hordelymus europaeus* (Poaceae: *Triticeae*). *Hereditas* 143, 103–112.
- Flaksberger, K.A. (1913) *Triticum monococcum* L. *Tr. Byuro prikl. bot.* 6(10), 669–695.
- Forsskål, P. (1775) *Flora Aegyptiaco-Arabica: Sive Descriptiones Plantarum, Quas per Aegyptum Inferiorem et Arabiam Felicem Detexit, Illustravit Petrus Forsskål, Prof. Haun.; Post Mortem Auctoris Edidit Carsten Niebuhr* (ed: Niebuhr, C.). Möller, København [Copenhagen].
- Frederiksen, S. (1986) Revision of *Taeniatherum* (Poaceae). *Nordic J. Bot.* 6, 389–397.
- Frederiksen, S. (1991a) Taxonomic studies in *Dasypyrum* (Poaceae). *Nordic J. Bot.* 11, 135–142.
- Frederiksen, S. (1991b) Taxonomic studies in *Eremopyrum* (Poaceae). *Nordic J. Bot.* 11, 271–285.
- Frederiksen, S. and Peterson, G. (1998) A taxonomic revision of *Secale* L. (*Triticeae*, Poaceae). *Nordic J. Bot.* 18, 399–420.
- Gandilyan, P.A. (1972) On the wild species of *Triticum* of the Armenian SSR. *Botan. zhurn.* 57, 173–181.
- Gould, F.W. (1947) Nomenclatural changes in *Elymus* with a key to the California species. *Madroño* 9, 120–128.
- Gould, F.W. (1949) Nomenclatural changes in Arizona grasses. *Madroño* 10, 94–95.
- Grossheim, A.A. (1939) *Flora Kavkaza*, vol. 1, 2nd edn. Trudy Botanicheskogo Instituta, Akademiia Nauk, SSSR Azerbaidzhanskii Filial. vol. 8. Izdatel'stvo Azfan, Baku, Azerbaijan.
- Hackel, E. (1887) *Gramineae*, vol. 2, part 2. Die natürlichen Pflanzenfamilien. Wilhelm Engelmann, Leipzig, Germany, pp. 1–97.
- Haussknecht, H.C. (1899) *Symbolae ad floram graecam: Aufzählung der im Sommer 1885 in Griechenland gesammelten Pflanzen*. Mitth. Thüring. Bot. Ver., n.s., 13, 18–77.
- Hitchcock, A.S. (1951) *Manual of Grasses of the United States*, 2nd edn. (rev: Chase, A.). U.S. D.A. Miscellaneous Publication No. 200. U.S. Government Printing Office, Washington, DC.
- Jensen, K.B. and Wang, R.R.-C. 1997. Cytological and molecular evidence for transferring *Elymus coreanus* from the genus *Elymus* to *Leymus* and molecular evidence for *Elymus californicus* (Poaceae: *Triticeae*). *Int. J. Plant Sci.* 158, 872–877.
- Körnicke, F.A. (1885) *Die Arten und Varietäten des Getreides*. Vol. 1 of Körnicke, F.A. and Werner, H. *Handbuch des Getreidebaues*. Paul Parey, Berlin.
- Lawrence, G.H.M. (1968) *Taxonomy of vascular plants*. The MacMillan Company, New York, USA.
- Linde-Laursen, I. and Seberg, O. (2001) Karyotypes of *Elymus scabrifolius* (Poaceae: *Triticeae*) from South America studied by banding techniques and *in situ* hybridization. *Hereditas* 135, 41–50.
- Link, J.H.F. (1834) *Symbolae ad floram graecam*. *Linnaea* 9, 129–132.
- Linnaeus, C. (1753) *Species Plantarum*. Impensis Laurentii Salvii, Stockholm.
- Liu, Q., Ge, S., Tang, H., Zhang, X., Zhu, G. and Lu, B.-R. (2006) Phylogenetic relationships in *Elymus* (Poaceae: *Triticeae*) based on the nuclear ribosomal internal transcribe spacer and chloroplast *trnL-F* sequences. *New Phytol.* 170, 411–420.
- Löve, Á. (1984) Conspectus of the *Triticeae*. *Feddes Repert.* 95, 425–521.

- Löve, Á. and Löve, D. (1961) Some nomenclatural changes in the European flora: II. Subspecific categories. *Bot. Not.* 114, 48–56.
- Mason-Gamer, R.J. (2001) Origin of North American *Elymus* (Poaceae: *Triticeae*) allotetraploids based on granule-bound starch synthase gene sequences. *Syst. Bot.* 26, 757–768.
- McNeill, J., Barrie, F.R., Burdet, H.M., Demoulin, V., Hawksworth, D.L., Marhold, K., Nicolson, D.H., Prado, J., Silva, P.C., Skog, J.E., Wiersema, J.H. and Turland, N.J. (2006) *International Code of Botanical Nomenclature (Vienna Code)*. Regnum Vegetabile 146. A. R.G. Gantner Verlag KG, Ruggell, Liechtenstein.
- Nelson, E.E. (1904) Some western species of *Agropyron*. *Bot. Gaz.* 38, 378.
- Nevski, S.A. (1934) Agrostological studies IV. The systematics of the tribe Hordeae. *Trudy Botanicheskogo Instituta Adakademii Nauk SSSR. Ser. 1, Flora I Sistematika Vysshikh Rastenii* 1, 9–32.
- Pavord, A. (2005) *The naming of names: the search for order in the world of plants*. Bloomsberry, New York, USA.
- Petersen, G., Seberg, O., Yde, M. and Berthelson, K. (2006) Phylogenetic relationships of *Triticum* and *Aegilops* and evidence for the origin of the **A**, **B**, and **D** genomes of common wheat (*Triticum aestivum*). *Molec. Phylog. Evol.* 39, 70–82.
- Pinto, R.J.B., Scapim, C.A., Alvarez, J.B., Martin, L.M.M. (2003) Comparative analysis of grain yield and breadmaking quality between hexaploid tritordeum lines and other winter cereals. *Bragantia* 62, 117.
- Pursh, F.T. (1813) *Flora America Septentrionalis: Or a Systematic Arrangement and Description of the Plants of North America; Containing, Besides What Have Been Described by Preceding Authors, Many New and Rare Species, Collected During Twelve Years Travels and Residence in That Country*. White, Cochrane, London.
- Refoufi, A., Jahier, J. and Esnault, M. (2001) Genome analysis of *Elytrigia pycnantha* and *Thinopyrum junceiforme* and their putative natural hybrid using the GISH technique. *Genome* 44, 708–715.
- Sallares, R. and Brown, T.A. (2004) Phylogenetic analysis of complete 5' external transcribed spaces of the 18s ribosomal RNA genes of diploid *Aegilops* and related species (*Triticeae*, Poaceae). *Genet. Res. Crop Evol.* 51, 701–712.
- Salomon, B. and Lu, B.-R. (1992) Genomic groups, morphology, and sectional delimitation in Eurasian *Elymus* (Poaceae, *Triticeae*). *Plant Syst. Evol.* 180, 1–13.
- Scribner, F.L. and Smith, J.G. (1897) Native and introduced species of the genera *Hordeum* and *Agropyron*. *Bull. Div. Agrostology, U.S.D.A.* 4, part 5, 23–36.
- Seberg, O., Frederiksen, S., Baden, C. and Linde-Laursen, I. (1991) *Peridictyon*, a new genus from the Balkan Peninsula, and its relationship with *Festucopsis* (Poaceae). *Willdenowia* 21, 87–104.
- Simon, B.K. (2007) *Grassworld: Interactive key and information system on world grasses*. *Kew Bull.* 65, 475–484.
- Slageren, M.W. van (1994) *Wild Wheats: A Monograph of Aegilops L. and Amblyopyrum (Jaub & Spach) Eig* (Poaceae). Wageningen Agricultural University, Wageningen, Netherlands.
- Soreng, R.J., Peterson, P.M., Davidse, G., Judziewicz, E., Zuloaga, F.O., Filgueiras, T.S. and Morrone, O. (2003) Catalogue of New World Grasses (Poaceae): IV. Subfamily Pooideae. *Contrib. from the U.S. Nat. Herbarium* 48, 1–730.
- Stewart, A.V., Ellison, N., Salomon, B. and Connor, H.E. (2005) Genomic constitution of New Zealand *Triticeae*. *Czech J. Genet. Plant Breed.* 41(Special Issue), 98–100.
- Sun, G.-L., Tang, H. and Salomon, B. (2006) Molecular diversity and relationships of North American *Elymus trachycaulus* and the Eurasian *E. caninus* species. *Genetica* 127, 55–64.
- Svitashev, S., Salomon, B., Bryngelsson, T. and Bothmer, R. von (1996) A study of 28 *Elymus* species using repetitive DNA sequences. *Genome* 30, 1093–1101.
- Thellung, A. (1918) Neuere Wege und Ziele der botanischen Systematik. *Naturwissenschaftliche Wochenschr.*, n.s., 17, 465–475.

- Trinius, C.B. von (1829) *Flora altaica*, vol. 1. G. Reimerifor, Berolini [Berlin].
- Tsvelev, N.N. (1976) *Zlaki SSSR* [Grasses of the Union of Soviet Socialist Republics]. Nauka, Leningrad [St. Petersburg], U.S.S.R. [Russia].
- Tutin, T.G., Heywood, V.H., Burges, N.A., Moore, D.M., Valentine, D.H., Walters, S.M. and Webb, D.A. (Eds.) (1980) *Flora Europaea*, Vol. 5. Cambridge University Press, Cambridge, England.
- Verushkine, S. and Shechurdine, A. (1933) Hybrids between wheat and couch grass: Fertile *Triticum-Agropyrum* hybrids of great scientific and practical interest. *J. Hered.* 24, 329–335.
- Wang, R.R.-C. and Jensen, K. (1994) Absence of **J** genome in *Leymus* species (Poaceae: *Triticeae*): Evidence from DNA hybridization and meiotic pairing. *Genome* 37, 231–235.
- Wheeler, D.J.B., Jacobs, S.W.J. and Whalley, R.D.B. (2002) *Grasses of New South Wales*, 3rd edn. University of New South Wales, Armidale, New South Wales, Australia.
- Wilson, F.D. (1963) Revision of *Sitanion* (*Triticeae*, Gramineae). *Brittonia* 15, 303–323.
- Wu, Z.-Y. and Raven, P.H. (Eds.) (2007) *Flora of China Illustrations*. vol. 22. Poaceae. Science Press, Beijing, and Missouri Botanical Garden Press, St. Louis.
- Wu, Z.-Y., Raven, P.H. and Hong, D.-Y. (Eds.) (2006) *Flora of China*. vol. 22. Poaceae. Science Press, Beijing, and Missouri Botanical Garden Press, St. Louis.
- Yamane, K. and Kawahara, T. (2005) Intra- and interspecific phylogenetic relationships among diploid *Triticum-Aegilops* species (Poaceae) based on base-pair substitutions, indels, and microsatellites in chloroplast noncoding sequences. *Amer. J. Bot.* 92, 1887–1898.
- Yan, J. and Yang, J.-L. (1999) *Xiao mai zu sheng wu xi tong xue*. vol. 3. Beijing Shi, Beijing, P. R.C.
- Yen, C., Yang, J.-L. and Yang, Y. (2005a) Hitoshi Kihara, Áskell Löve and the modern genetic concept of the genera in the tribe *Triticeae* (Poaceae). *Acta Phytotaxon. Sin.* 43, 82–93.
- Yen, J.-I., Yang, J.-L. and Baum, B.R. (2005b) *Douglasdeweya*: A new genus, with a new species and a new combination (*Triticeae*: Poaceae). *Can. J. Bot.* 83, 413–419.
- Zhang, H.B. and Dvůrák, J. (1991) The genome origin of tetraploid species of *Leymus* (Poaceae: *Triticeae*) inferred from variation in repeated nucleotide sequences. *Am. J. Bot.* 78, 871–884.

Chapter 2

Triticeae Genetic Resources in *ex situ* Genebank Collections

Helmut Knüpfper

Abstract Of the ca. 350 species and ca. 30 genera estimated for the *Triticeae*, 111 species of 19 genera are either cultivated or useful wild species. These species are listed together with their main uses in the first part of this chapter. The second part provides an overview of *Triticeae* holdings in the world's genebanks. Starting from a survey of available online information resources, 1,278,000 accessions of *Triticeae*, belonging to 35 genera (among them 12 hybrid genera) and almost 300 species and, comprising one-fifth of the estimated world germplasm holdings, are documented. The survey includes data from 295 genebanks. Summaries of the world holdings per genus, species, regions where the genebanks are located, and the largest collections of the major genera are provided. For the larger genera, i.e., *Triticum*, *Hordeum*, \times *Triticosecale*, *Aegilops*, *Secale*, *Elymus* and *Agropyron*, the worldwide germplasm collections are surveyed with more details. Existing international or regional cooperation programmes, germplasm databases and cultivar registers with pedigree information for these genera are briefly described. For the major cereal crops, core collections and genetic stocks collections are also mentioned. Due to its growing importance as a model plant in genomics research, the genus *Brachypodium* closely related to the *Triticeae* is also included in the surveys. The paper aims at providing background information for plant breeders and crop plant researchers on the germplasm available in *ex situ* genebank collections, to make this wealth of material more easily accessible.

2.1 Introduction

The grass tribe *Triticeae* includes some of the major cereal crop species of the world, namely *Triticum aestivum* (bread wheat), *durum* (durum wheat), *Secale cereale* (rye), and *Hordeum vulgare* (barley), the modern cereal \times *Triticosecale*

H. Knüpfper (✉)

Genebank Department, Leibniz Institute of Plant Genetics and Crop
Plant Research (IPK), Corrensstrasse 3, D-06466 Gatersleben, Germany
e-mail: knupffer@ipk-gatersleben.de

(triticale), plus about 350 other species (Löve 1984) (cf. also Chapter 1). Besides the cereals, species of various genera of the *Triticeae* are used, often cultivated, for animal forage and fodder, for revegetation and soil erosion control (Hanelt 2001). Several species are considered as weeds, some of them with high colonising potential (USDA-ARS 2008).

Species belonging to various *Triticeae* genera have been used, or have the potential of being used, in plant breeding and breeding research of the major cereals, particularly wheat, due to a high rate of crossability. These species and their uses are further described in Section 2.3. Wild relatives of crops, together with the landraces, traditional and modern cultivars, and genetic stocks form the genetic resources of crops in the *Triticeae*. The whole tribe can be considered as genetic resources for the major cereals, but also for species used for other purposes mentioned above.

The need to safeguard existing landraces and wild relatives of crop species as a resource for plant breeders was first recognized by Vavilov in the beginning of the 20th century, who established collections of cultivated plants and their wild relatives, now known as genebanks. Many national, regional and institutional genebanks have been established since then, particularly in the early 1970s. In 1996, the Food and Agriculture Organization (FAO) concluded that there were more than 1300 genebanks conserving over six million accessions *ex situ* (FAO 1996). In addition, more than half a million accessions were being held in field genebanks (Hintum and Menting 2003).

It is often difficult for breeders and researchers to locate relevant information and to identify potential sources of plant material for their studies. This chapter aims at giving an overview of *ex situ* collections holding plant material of *Triticeae* genera and species, with a focus on genebanks. Many wild species are also kept in botanical gardens, but these are more difficult to include in such a review. Here, we provide an updated and more extensive review of the *Triticeae* genetic resources than previous studies (e.g., Bothmer et al. 1992) with an exhaustive list of the cultivated species (Section 2.3) and an overview of existing genetic resources collections, maintaining living, mostly seed material of *Triticeae* species (Section 2.4).

2.2 Material and Methods

This section describes the sources of information used in the present study, and the methods of processing, standardising and summarising the information.

2.2.1 Information Sources: Online Databases and Reports

Information about cultivated *Triticeae* species (Section 2.3) was extracted from “Mansfeld’s Encyclopedia of Agricultural and Horticultural Crops” (Hanelt 2001), which is also available online (Mansfeld Database, 2008). The resulting list of species and their uses was complemented by information available online

from the World Economic Plants Report from the United States Department of Agriculture (USDA-ARS 2008).

For the overview of *Triticeae* collections in genebanks (Section 2.4), an initial inventory of germplasm collections was derived primarily from two data sources, namely, (1) the “Bioversity Directory of Germplasm Collections” (Bioversity 2007), maintained by Bioversity International, and (2) the FAO’s “World Information and Early Warning System” (WIEWS) (FAO 2007), which was created initially to collate summary information about germplasm collections worldwide for the FAO Conference on Plant Genetic Resources in 1996 (FAO 1996, 1997). Both information systems started from the same data collection effort in the mid 1990s, but were apparently updated independently at different times till 2004.¹ The overlap between the two data sources is still considerable, but each data set has information about genebanks not included in the other one. I have attempted to bring the information from the two data sets together, removing duplicate information, to obtain a more complete picture than by consulting only one of them. A number of genebanks reported to contain *Triticeae* taxa do not provide the number of accessions. These genebanks are included in the present study, but with numbers of accessions equalling zero. In some cases it was possible to use other information sources to get the actual number of accessions, but 20 of the 295 genebanks analysed still contribute no accessions to the total.

To make the present survey as complete as possible and use up-to-date information, additional data sets were downloaded from the System-wide Information Network for Genetic Resources (SINGER), the federated database of germplasm accessions of the International Agricultural Research Centres (IARCs) of the Consultative Group of International Agricultural Research (CGIAR) (SINGER 2008), and the European Plant Genetic Resources Search Catalogue (EURISCO 2008). For the European region, information from the “Central Crop Databases” of the European Cooperative Programme for Plant Genetic Resources (ECPGR) and recent meeting reports (e.g. Maggioni et al. 2008) was also included. In addition, online databases of some major genetic resources collections were searched, and genebank curators were contacted to request summary information about their *Triticeae* holdings. Finally, the “Crop Strategies” for wheat (including also rye, triticale and *Aegilops*) and barley, developed by the Global Crop Diversity Trust, were also taken into account.

Major online information sources used are tabulated in the Appendix.

2.2.2 Information Extraction and Processing

To collect the information about *Triticeae* holdings from the above-mentioned Bioversity and FAO databases, lists of all scientific names of each of these two

¹ The situation of autumn 2007 described here changed for the FAO WIEWS database meanwhile considerably. During the past seven years, FAO has been constantly updating the information in WIEWS, using the best available sources, either direct information from the respective genebanks or combined information systems such as EURISCO (S. Diulgheroff, pers. comm. March 2009).

sources were retrieved from the respective websites. Since neither of the two databases provide the possibility of downloading query results, series of queries were submitted for each generic name belonging to the *Triticeae* (including synonyms). The resulting information about holding institutions, scientific names, numbers of accessions, and information source/date of last update, was copied into text files and parsed to extract the relevant information into a database.

Information about institutions holding *Triticeae* germplasm from the other sources consulted replaced information from Bioversity/FAO in cases where it was of more recent origin.

In the FAO WIEWS database, institutions are identified by so-called FAO Institution Codes consisting of three letters for the country (ISO codes) and three digits. The Bioversity Directory does not use these codes. Many other databases consulted, as well as the two Crop Strategies, are also referring to these same FAO codes, thus allowing matching information related to the same institution.

2.2.3 Handling of Nomenclature

In any attempt to bring together information from numerous sources about germplasm holdings, the problem of harmonising or standardising scientific nomenclature between the data sources arises. This is true for both taxon-level data sources (for example, when compiling a continental flora from numerous country floras that use different taxonomy and nomenclature), and accession-level data sources (e.g., in the development of crop-specific databases from accession data provided by individual genebanks using different classification systems) and is due to the use of different nomenclatures in different collections.

The two meta-databases used (Bioversity 2007; FAO 2007) apparently took unaltered scientific names from their contributing genebanks, so that they reflect the names employed in the individual genebanks. In addition, names are reported with different level of details, from genus only (e.g., “*Triticum* sp.”) to detailed infraspecific classifications (e.g., “*Triticum dicoccon* subsp. *transcaucasicum* var. *nivicum*”). Consequently, to summarise the information about genebank holdings by genus or species, it was necessary to group apparent synonyms together, and to relate them to a single accepted name.

First, names not conforming to the rules of botanical nomenclature (McNeill et al. 2006) were transformed by guessing what they meant. For example, in many *Triticum* and *Hordeum* names, only the generic name was given, followed by one or several botanical varietal epithets but without a specific epithet. Fortunately, in most cases, it was clear which species was involved. For instance, the name “*Triticum albidum, ferrugineum, milturum, lutescens, subferrugineum*” apparently refers to a mixture (population) of botanical varieties of *Triticum aestivum*.

To weed out spelling errors and identify possible synonyms at the generic and species levels, the “Taxonomic Nomenclature Checker” (TNC-GRIN 2008) was used. It allows, in a stepwise process, to crosscheck large lists of scientific plant names against the names accepted in the Genetic Resources Information Network of the USDA. For misspellings, it offers alternative correct spellings using a string-matching algorithm, and for synonyms, it points to the corresponding accepted names. The check was carried out using the following steps:

1. checking generic names for spelling errors, correcting the erroneous ones;
2. checking binomials (genus + species) for spelling errors (i.e., for a misspelled species epithet in a valid genus), correcting the erroneous ones;
3. checking for synonyms against GRIN-Taxonomy, transferring synonymic names to their accepted names (according to GRIN-Taxonomy);
4. finally, the resulting list was checked for names without a specific epithet. If a possible species epithet could be inferred from the “original scientific name”, it was added. Various Internet sources were used, usually trying Google and looking for “serious” nomenclatural websites. Otherwise, the abbreviation “sp.” was used.

The resulting list of species and some infraspecific taxa was then compared with the treatment of the *Triticeae* in the recent volume of “Flora of North America” (Barkworth et al. 2007), in consultation with M. Barkworth and R. von Bothmer, the authors of Chapter 1 “Scientific names in the *Triticeae*”. For a few genera, other monographic treatments were taken as basis, i.e., Bothmer et al. (1995) for *Hordeum*, the Russian “Wheat Flora” (Dorofeev et al. 1979) for *Triticum*, Hammer (1980) for *Aegilops*, and Hammer et al. (1987) for *Secale*.

2.3 List of Cultivated and Useful *Triticeae* Species

This chapter provides a list of species and their main uses. It is based mainly on Mansfeld’s Encyclopedia (Hanelt 2001) and online information from USDA-ARS (2008), as well as references cited in these sources. Additional economic uses of *Triticeae* species are reported by Tzvelev (1976). Besides the uses listed below, many species are also characterised as weeds (cf. USDA-ARS 2008), but these are not included here. Of the 27 genera of *Triticeae* recognised by Barkworth and Bothmer (2008; Chapter 1), nineteen genera (including three hybrid genera) and 111 species are listed. Due to its importance for genomics research, the closely related genus *Brachypodium* not included in the *Triticeae* by Barkworth (pers. comm.) is also treated below.

2.3.1 *Aegilops* – Goat Grass

Species of *Aegilops* played a central role in the evolution of tetra- and hexaploid wheat taxa as donors of the genomes **B** (probably from the *Sitopsis* section *Ae. longissima*, *Ae. searsii*, *Ae. speltoides*) and **D** (*Ae. tauschii*, probably its

subspecies *strangulata*). They are also important sources of new genes and alleles in wheat breeding, especially for the transfer of resistance against fungal diseases (e.g., *Ae. umbellulata* – against leaf rust; *Ae. peregrina* – against powdery mildew; *Ae. comosa* – against yellow rust and others; *Ae. caudata* – against stem rust; *Ae. speltoides* – against leaf rust; *Ae. ventricosa* – against eyespot (Dosba and Dossinault 1973)). For example, mildew-resistant wheat–*Ae. markgrafii* introgression lines have been successfully utilised in breeding programmes with high-yielding wheat cultivars (Schubert 2000). *Aegilops* spp. are also useful donors for pest resistance (*Ae. cylindrica*, *Ae. geniculata*, *Ae. markgrafii*, *Ae. neglecta*, *Ae. tauschii*, *Ae. ventricosa*), drought (*Ae. longissima*) and salt tolerance (*Ae. ventricosa*), and as genetic resources, i.e., potential gene sources (*Ae. bicornis*, *Ae. biuncialis*, *Ae. columnaris*, *Ae. comosa*, *Ae. crassa*, *Ae. juvenalis*, *Ae. kotschyi*, *Ae. searsii*, *Ae. sharonensis*, *Ae. triuncialis*, *Ae. umbellulata*, *Ae. uniaristata*, and *Ae. vavilovii*). For *Ae. speltoides*, the use as animal food is also reported (Kimber and Feldman 1987), e.g. from Iraq (Bor 1968). According to Tzvelev (1976), *Ae. cylindrica*, *Ae. tauschii* and *Ae. crassa* are good forage grasses consumed by all species of livestock.

2.3.2 × *Aegilotriticum*

Natural and artificial hybrids between *Aegilops* and *Triticum* species; some have potential value for breeding (e.g., Kimber and Sears 1987).

2.3.3 *Agropyron* – *Wheatgrass*

According to Tzvelev (1976), all species of *Agropyron* occurring in the former USSR are valuable forage crops with good drought tolerance.

A. cristatum subsp. *cristatum* is a forage grass for pasture and hay (mostly Southern Russia, North America), drought-resistant and tolerant to low temperatures, and is also used as crossing partner with *Elymus* spp. in grass breeding programmes. Subsp. *pectinatum* has been experimentally cultivated in southern parts of the former Soviet Union.

A. desertorum is an important range grass in North America, used for hay, pasture and erosion control. It is long-lived, drought- and frost-resistant, and also cultivated in the dry steppe belt of the former Soviet Union.

A. fragile is sown and used as range grass and for revegetation in western North America and cultivated in southern regions of the former Soviet Union as forage grass.

According to USDA-ARS (2008), *A. michnoi* and *A. mongolicum* are used as animal food (forage, fodder). Tzvelev (1976) reports the use of *A. dasyanthum*, *A. tanaiticum*, *A. cimmericum*, *A. michnoi* and *A. fragile* for fixing moving sand dunes.

There is disagreement over the treatment of the species in *Agropyron*, Barkworth et al. (2007) arguing that the genetic structure of accessions being used in North America has been modified to such an extent that application of names used for natural populations is inappropriate.

2.3.4 *Amblyopyrum*

A. muticum has potential as genetic resource for wheat breeding (USDA-ARS 2008).

2.3.5 *Brachypodium* – *False Brome*

This genus is sometimes included in *Triticeae* (e.g., Hanelt 2001), but also often placed in a separate tribe, *Brachypodieae*. However, since it contains an important model species for genomics research, it is also mentioned here.

B. distachyon has become an important model species for functional genomics studies in temperate grasses and cereals (e.g., Draper et al. 2001; Hasterok et al. 2004; Garvin et al. 2008), for which it has many suitable qualities, including a small genome (~300 Mbp), diploid and polyploid accessions, a small physical stature, self-fertility, a short lifecycle, simple growth requirements, and an efficient transformation system (from International Brachypodium Initiative 2008).

Breeding work on *B. pinnatum* for developing forage cultivars has been reported from Italy (Ceccarelli et al. 1975). The species is also used for ornamental purposes (USDA-ARS 2008),

2.3.6 *Dasypyrum* – *Mosquitograss*

D. villosum has been used as crossing partner in wheat breeding programmes (e.g. Pace et al. 1990), and as animal forage (Tzvelev 1976).

2.3.7 *Elymus* – *Wheatgrass, Wild Rye*

E. canadensis is a pasture grass, occasionally cultivated in its native area (North America); it is also used for restoration in disturbed prairie sites and is a crossing partner of other *Elymus* spp. in grass breeding programmes.

E. caninus is grown for forage in Russia and North China.

E. dahuricus is grown in grass mixtures to increase the productivity of slowly establishing long term pastures in North America. It is a short rotation hay and pasture crop.

E. elymoides is used for erosion control and reclamation of oil shale and coal mining areas in North America, as well as for animal food; breeding lines for experimental cultivation have been developed.

E. fibrosus is a minor forage crop in Russia.

E. glaucus is used as grass cover in windbreaks, shrub plantings, etc., and has potential value as pasture and forage crop.

E. hoffmannii is reported as a forage crop.

E. hystrix is cultivated for erosion control and for ornamental purposes (Tzvelev 1976).

E. lanceolatus is cultivated as forage grass and pasture crop (forage type) as well as for soil stabilization of disturbed sites (ground-cover type). It is also a crossing partner in forage grass breeding programmes.

E. macrourus is used for soil fixation and revegetation of mining stockpiles (Tzvelev 1976).

E. mutabilis is a frost-resistant forage grass cultivated in Sakha (Yakutia).

E. nutans has been successfully tested as drought-tolerant forage crop in the former Soviet Union (Tzvelev 1976).

E. repens (earlier often treated in other genera, e.g., *Elytrigia* and *Agropyron*; see Chapter 1) is an aggressive weed in temperate regions, but sometimes sown against soil erosion and, in Scandinavia, as forage plant for hay. In Germany, it is a minor medicinal crop. It is also used in trials to obtain perennial wheat cultivars and in forage grass breeding programmes for hybridizations with other *Elymus* and *Agropyron* spp. The North American cultivar 'Newhy' selected from *E. repens* × *Pseudoroegneria spicata* seems promising as forage grass for moderately saline and semiarid rangeland. The cultivar reflects selection for the bunching growth habit of *Pseudoroegneria* rather than the aggressively rhizomatous growth habit of *E. repens*.

E. semicostatus has been tried as drought-resistant pasture grass in North America and Japan.

E. sibiricus is rarely cultivated as forage grass in Western Siberia and the Amur region, and formerly also in North America.

E. trachycaulus is cultivated for forage and pasture in grass mixtures, also for soil conservation and reclamation in South Russia, Siberia and North America. It has been the first native American grass species widely used for reseeding of range land.

E. villosus is used for erosion control.

E. virginicus is planted for restoring native vegetation in disturbed prairie sites, and used as forage.

2.3.8 *Eremopyrum* – False Wheatgrass

E. orientale and other species of the genus are used as a pasture crop with good performance in moist years, and for hay as animal food (Tzvelev 1976).

2.3.9 *Heteranthelium*

H. piliferum is used as pasture crop (Tzvelev 1976).

2.3.10 *Hordeum* – *Barley*

H. bogdanii is reported as pasture grass (Tzvelev 1976).

H. brevisubulatum subsp. *violaceum* used to be cultivated as forage grass in the European part of former Soviet Union.

H. bulbosum used to be grown as pasture and forage grass in the USA, Argentina and Australia, and Tzvelev (1976) recommended its cultivation under irrigation in dry areas of the former Soviet Union. The roots can be eaten (Tanaka 1976). It has also been used for the production of doubled haploid lines in barley and wheat through chromosome elimination. *H. bulbosum* has a potential as gene donor of, particularly, disease resistance for barley and is used in barley breeding research and breeding programmes in Germany (Szigat et al. 1998) and New Zealand.

H. capense is reported as genetic resource for barley breeding (USDA-ARS 2008).

H. jubatum is widely cultivated as an ornamental grass for dry bouquets (Tzvelev 1976).

H. vulgare subsp. *spontaneum* is the progenitor of cultivated barley, completely cross-compatible, and a useful resource for barley breeding. It is also a good forage grass (Tzvelev 1976).

H. vulgare subsp. *vulgare*, the cultivated barley, is one of the main cereals grown worldwide. It is a short- and cool-season crop, grown as spring or winter form. It tends to be more salt-tolerant than wheat. It is used for animal feed, main cereal for malt production, e.g., for beer brewing and whisky production, and an important human food. It is also a favoured model organism in biological, particularly genetics and genomics research in the *Triticeae*.

2.3.11 *Kengyilia*

K. grandiglumis is used as animal food (forage).

2.3.12 *Leymus* – *Wildrye*

According to Tzvelev (1976), several species of this genus are used for pasture and hay despite of their relatively low nutritional value, especially in desert and steppe environments and on sandy soils. Many species are salt and drought tolerant. The straw of some of the larger species can be used for the production of paper.

L. angustus is grown as pasture grass in North America; a valuable autumn and winter grazing, salt tolerant, and most productive on clay-loam and clay soils. It is used in a hybridization breeding programme with other taxa of *Leymus*.

L. arenarius is grown for stabilization of coastal sand dunes, sometimes also in inland sites in North and Western Europe, Greenland and North America. Until the beginning of the 20th century, its relatively large grains were collected for human food in Iceland. It is still used as a fodder grass.

L. cinereus supplements *Thinopyrum ponticum* in pasture programmes for saline and alkaline soils in Western USA. It provides early spring, late autumn and winter grazing for livestock and wildlife, is used to reseed rangelands, mine spoils and for general soil reclamation. It is a valuable thermal protection of livestock and a windbreak. It is also used in cross-breeding programmes with other *Leymus* spp.

L. mollis is planted for stabilization of coastal sand dunes in Japan, USA, and formerly in Iceland. It is used for revegetation and reclamation of disturbed sandy coastal sites in Alaska, where it is also a traditional forage grass and used for thatching and basketry.

L. racemosus is planted for stabilization of moving inland sand dunes in deserts in Eurasia, and formerly in Western USA. Grains from wild stands were collected for human food by nomads in Central Asia. It is also used for forage.

L. triticoides is planted for soil stabilization along roadsides, channels, river slopes and as forage grass and for amelioration of poorly drained alkaline soils in Western USA.

Several additional *Leymus* species are reported as being used for forage (e.g., *L. condensatus*, *L. karelinii*, *L. secalinus*, *L. tianschanicus*) and fodder (e.g., *L. chinensis*, *L. divaricatus*, *L. multicaulis*, and *L. ramosus*) (USDA-ARS 2008).

2.3.13 *Pascopyrum* – *Wheatgrass*

P. smithii is a range grass in natural pastures. It is cultivated for hay, pasture, erosion control, revegetation of disturbed lands and reseeded of rangelands in North America. It shows promise for reclamation of saline soils, but is a poor seed producer.

2.3.14 *Psathyrostachys* – *Wildrye*

P. juncea is grown as pasture plant in arid zones of former Soviet Union and in the prairie belt of North America. Breeding programmes in Siberia and USA were established in the first half of the 20th century; recently induced tetraploid strains have been used as breeding material. It is also considered a genetic resource for disease resistance in wheat (e.g., Chen et al. 1988; Plourde et al. 1990). Other *Psathyrostachys* species are also reported as forage grasses (Tzvelev 1976).

2.3.15 *Pseudoroegneria* – *Wheatgrass*

P. spicata is grown for erosion control and revegetation, as well as forage. Its f. *inermis* is drought-resistant, grown as forage grass in Western USA, mainly for reseeding of rangelands. The f. *spicata* is a forage grass of natural pastures, sown in North America for reseeding and reestablishment of rangelands, recommended for soil reclamation, very drought-resistant and palatable. See also under *Elymus repens*.

2.3.16 *Secale* – *Rye*

S. cereale subsp. *afghanicum* is used in breeding of cultivated rye, subsp. *cereale*.

S. cereale subsp. *ancestrale* is the progenitor of the cultivated subsp. *cereale*, and a potential genetic resource for rye breeding.

S. cereale subsp. *cereale* is a major cereal with winter and spring forms, grown from North and East Europe to Siberia, to a lesser extent also in the remaining Europe, North America, Argentina, southern Africa, more seldom in the Near East, North Africa, South America and Australia. It is the most cold-tolerant cereal; it also tolerates dry hot summers and acid soils. It is mainly used for (black) bread, crisp bread, starch, animal fodder, and for alcoholic beverages; as forage plant mainly in Europe, for soil conservation in USA and Australia, as host plant for ergot (*Claviceps purpurea*) for pharmaceutical uses, e.g., in India and Europe. The straw is used for thatching. Wheat-rye translocations derived from rye or 8× triticale have been widely used in wheat breeding to improve productivity, adaptiveness, as well as disease and insect tolerance (Rabinovich 1998).

S. cereale var. *multicaule* is a perennial rye, formerly grown in Eastern and Central Europe in forest clearings or as component of shifting cultivation for forage (first year of cultivation) and grain (following seasons).

S. cereale subsp. *dighoricum* and subsp. *segetale* are potential genetic resources for rye breeding.

S. ×derzhavinii, an artificial hybrid between *S. strictum* and *S. cereale*, is reported to be introduced into cultivation as perennial rye.

S. strictum is possibly the progenitor of the annual *Secale* taxa.

S. strictum subsp. *africanum* is used for food (cereal) and as genetic resource for rye breeding.

Wild stands of *S. strictum* ssp. *strictum* provide a good pasture and hay (Tzvelev 1976). Armenian populations of this subspecies have been interpreted as relics of former cultivation of perennial rye (as *S. daralagesi*); introgressions to sympatric cultivated rye have been observed. In USA it has been tried as green manure, cover crop, forage grass. Nowadays it is recommended in Western USA for erosion control, early spring forage and as a cover crop. Earlier reports indicate that it has been used for winter pasture in regions with winter rainfall.

S. vavilovii is a genetic resource for rye breeding.

2.3.17 *Thinopyrum* – *Wheatgrass*

T. elongatum is a good forage plant with high salt tolerance and is used in wide crosses for wheat breeding, along with other species of the genus (Tzvelev 1976).

T. intermedium subsp. *barbulatum* is grown in North America for hay and pasture. Subsp. *intermedium* is very salt tolerant. It is grown as forage crop for pasture and hay, for erosion control and soil reclamation, in North America and the steppe zone of Eastern Europe, often in mixture with alfalfa. Grains from wild plants have been harvested for human consumption in Southwest Asia. It is also used in wheat breeding programmes for transfer of disease resistances and in trials to obtain perennial wheat cultivars.

T. junceiforme is cultivated for stabilizing coastal sand dunes in Europe and Israel.

T. junceum is suitable for sand dune fixation, especially near the sea coast (Tzvelev 1976).

T. ponticum is a hay and pasture crop, and used for improvement of saline and alkaline soils in North America. It is high-yielding but not very palatable.

2.3.18 \times *Triticosecale* – *Triticale*

Triticale is a man-made cereal (e.g. Müntzing 1974) derived from hybrids between 4 \times and 6 \times wheats and rye with subsequent chromosome doubling. It combines tolerance to diseases, drought and poor soils and the winter hardiness of rye with superior yield and grain quality of wheat. It is commercially grown in more than 30 countries, primarily for forage and feed, but also for unleavened bakery products.

2.3.19 *Triticum* – *Wheat*

T. aestivum is the most important cereal for bread, cultivated between 67°N and 45°S worldwide. Its grains, flour and straw are used for both human consumption and as forage.

T. aethiopicum is a traditional wheat crop in Ethiopia and the Southern Arabian Peninsula, adapted to higher rainfall regions.

T. araraticum, the ancestor of the cultivated *T. timopheevii*, evolved presumably between East Turkey, North Iraq and Transcaucasia from spontaneous hybridizations between *T. urartu* and a member of *Aegilops*, probably *Ae. speltoides*.

T. boeoticum (from Southeastern Turkey) is the progenitor of the cultivated einkorn and donor of the modified genome **A^m**, which characterizes *T. monococcum* and the third genome of *T. zhukovskyi*.

T. dicoccoides is the wild form and immediate progenitor of cultivated emmer. It is used in breeding programmes as crossing partner for its high grain protein content and resistance against fungal diseases.

T. dicoccon was once widely cultivated for bread and emmer flour products, from Southwest and Middle Asia, India to the Mediterranean region and Northern Europe, and later also in North America. Nowadays it is a relic crop, often mixed with other wheats. It is mainly used for porridge and the preparation of special sausages, seldom as animal feed.

T. durum is the second most important wheat species, especially in warmer and drier climates. It is used for bread. Its low gluten flour makes it especially suitable for pasta and similar products.

T. ispahanicum is a potential genetic resource, related to emmer.

T. karamyshevii is an endemic wheat species from Western Georgia, grown mostly as admixture of *T. macha*. It is important for wheat breeding because of its high grain protein content and low susceptibility to fungal diseases.

T. macha is a traditional Western Georgian endemic wheat; only winter types are known.

T. monococcum was formerly widely cultivated for bread, but it is now grown mainly for porridge and other forms of cooking, from Near East to the Iberian Peninsula, and from North Africa to Sweden. Nowadays it is sporadically grown as a relic crop. In Western Georgia it was the main component of the traditional “zanduri wheat” complex, grown together with *T. timopheevii* and *T. zhukovskyi* under humid mountain conditions. Nowadays it is grown mostly as grain forage, rarely for cracked grain foodstuff.

T. persicum is an endemic wheat from the Caucasus region, grown in pure stands or mixed with emmer, durum or soft wheat. The grains have high protein content, but low baking quality. The plants are mildew-resistant.

T. polonicum is a minor crop, usually an attendant of other wheats (mainly of durum or – in Ethiopia – Ethiopian wheat), in the region from the Mediterranean to China and Tibet.

T. sinskajae is a spontaneous naked mutant from an accession collected in Northern Turkey; a germplasm strain of it was officially released in Italy.

T. spelta was formerly rather widely cultivated for bread in parts of Europe, traditionally also in Southwest Asia and oases of North Africa. Nowadays it is mostly abandoned, being grown only as an admixture in common wheat. Recently it has become fashionable (e.g., Perrino et al. 1996; Cubadda and Marconi 1996).

T. sphaerococcum is a relic cereal crop in Northwest India and Pakistan.

T. timopheevii was one of the endemic Georgian “zanduri wheats” (cf. under *T. monococcum*), but it has completely disappeared from cultivation. It is important for wheat breeding because of its excellent resistance against major fungal diseases and pest insects; also for cytoplasmatic male-sterile forms and restorer genes.

T. turanicum is presently grown mostly as admixture in durum wheat fields, rarely in pure sowings; it was formerly more widespread in Near East countries,

the Levant, North Africa, and Middle Asia to Pakistan and Northwest India. It is resistant against fungal diseases.

T. turgidum has been cultivated since prehistoric times, temporarily also introduced to USA and Australia.

T. urartu, formerly often not distinguished from *T. boeoticum*, is the donor of genome A for most of the other species (*T. aestivum* in particular) of the genus (besides *T. monococcum*).

T. vavilovii is a genetic resource for wheat breeding.

T. zhukovskyi used to be a minor component of the endemic “zanduri wheat” complex in West Georgia (cf. under *T. monococcum*). Nowadays it has disappeared from cultivation. It is a hexaploid with the genome **GGAAmAm**, presumably resulting from a quite recent spontaneous hybridization between the diploid *T. monococcum* and the tetraploid *T. timopheevii* of the “zanduri” complex.

2.3.20 × *Tritordeum*

Plants of this hybrid genus between species of *Triticum* and *Hordeum* are a genetic resource for wheat breeding. Amphiploids between *H. chilense* and *Triticum* spp. have the potential to become a new crop when sufficient breeding effort is applied (Martín et al. 1996).

2.4 Overview of *ex situ* Collections of *Triticeae*

2.4.1 Overview by Countries and Institutions

The present survey includes 295 *Triticeae*-holding genebanks in 83 countries out of the ca. 1300 genebanks existing worldwide. The total number of *Triticeae* accessions is over 1,284,000 (cf. Table 2.1), thus making up about one-fifth of the world's germplasm holdings. Of these, 41.1% are held in 175 collections in Europe, 15.7% in 22 collections in North America, 14.9% in 31 collections in Asia, 7.8% in 38 collections in South America, 4.1% in 6 collections in Australia and New Zealand, and 3.7% in 19 collections in Africa. About 12.6% are maintained in four genebanks of CGIAR centres (Table 2.2). The largest genebank collections of *Triticeae* are listed in Table 2.3.

The genebank material documented in this survey belongs to 35 genera, among them 12 hybrid genera (cf. Table 2.1). The number of species names reported amounts to ca. 290, thus approaching the total number of *Triticeae* species as estimated by Löve (1984), i.e., ca. 350, rather closely (cf. Chapter 1).

According to a survey of FAO (FAO 1996, p. 92, Table 3.1), wheat, barley, triticale and rye are well represented in genebanks, with wheat and barley

Table 2.1 Total numbers of accessions of *Triticeae* genera. Column 2: Number of accessions named only to genus. Column 3: Number of accessions named to the level of species. The last column provides the number of holding institutions for each genus

Genus	Accs. to genus	Accs. to species	Total accs.	Genebanks
<i>Aegilops</i>	10340	21050	31401	61
× <i>Aegilotriticum</i>	568		568	8
<i>Agropyron</i>	1600	1510	3110	47
× <i>Agrotriticum</i>	41	0	41	4
<i>Amblyopyrum</i>	0	138	138	13
<i>Crithopsis</i>	0	9	9	5
<i>Dasypyrum</i>	99	510	609	21
× <i>Elyhordeum</i>	8	0	8	2
× <i>Elymotriticum</i>	5	0	5	2
<i>Elymus</i>	2585	1547	4132	46
× <i>Elytricum</i>	407	0	407	1
<i>Elytrigia</i>	278	29	307	12
× <i>Elytriticale</i>	101	0	101	1
× <i>Elytritolops</i>	6	0	6	1
<i>Eremopyrum</i>	73	145	218	9
<i>Festucopsis</i>	3	0	3	1
× <i>Haynaldoticum</i>	6	8	14	2
<i>Henrardia</i>	5	7	12	2
<i>Heteranthelium</i>	6	11	17	5
× <i>Hordale</i>	1	0	1	1
<i>Hordelymus</i>	0	58	58	11
<i>Hordeum</i>	63511	390097	453602	204
<i>Hystrix</i>	14	0	14	1
<i>Kengyilia</i>	0	33	33	4
<i>Leymus</i>	230	1061	1291	27
<i>Pascopyrum</i>	0	94	94	13
<i>Psathyrostachys</i>	97	315	412	21
<i>Pseudoroegneria</i>	114	296	410	12
× <i>Pseudoroelymus</i>	1	0	1	1
<i>Secale</i>	2069	19475	21550	94
<i>Taeniatherum</i>	63	131	194	16
<i>Thinopyrum</i>	24	716	740	33
× <i>Triticosecale</i>	31867	0	31867	71
<i>Triticum</i>	167133	565129	732262	223
× <i>Tritordeum</i>	28	0	28	3
<i>Triticeae</i> (not specified)	568	0	568	3
Total	281851	981319	1284231	295

having 784,500 and 485,000 accessions, respectively, thus amounting to 13% and 8%, respectively, of the estimated six million accessions worldwide. These figures differ from our findings (Table 2.2). Wheat and barley are followed by *Oryza* (rice, 7%), *Zea* (maize, 5%), *Phaseolus* (common bean, 4%), *Glycine* (soybean) and *Sorghum* (sorghum) (3% each).

Table 2.2 *Triticeae* holdings by continents. The CGIAR centres are grouped separately. The last row shows the figures given in the FAO report (FAO 1996, p. 465 ff.) for comparison. The figures for wheat and barley differ slightly between FAO 1996, p. 92 and pp. 465 ff. Abbreviations: *Trit.* – *Triticum*, *Hord.* – *Hordeum*, \times *Tr.* – \times *Triticosecale*, *Aeg.* – *Aegilops*, *Sec.* – *Secale*, *Ely.* – *Elymus*, *Agr.* – *Agropyron*, *Oth.* – other genera

Continent	Total	<i>Trit.</i>	<i>Hord.</i>	\times <i>Tr.</i>	<i>Aeg.</i>	<i>Sec.</i>	<i>Ely.</i>	<i>Agr.</i>	<i>Oth.</i>
Africa	47978	24154	23072	210	15	479	6	27	101
Asia	191565	120967	61161	241	6535	1354	189	928	206
Australia, Oceania	52383	25821	25389	0	740	111	3	298	21
Europe	528260	298650	175054	20317	12990	15845	2887	547	5208
North America	201787	89636	92794	4138	6218	3570	924	1227	4755
South America	99834	58774	38841	1066	632	148	120	69	191
CGIAR	162424	114260	37291	5895	4271	31	3	14	673
Total	1284231	732262	453602	31867	31401	21538	4132	3110	11155
cf. FAO 1996		788654	486724	40131	27132	21360	2665		

2.4.2 Overviews by Genera and Species

For some genera with relatively large numbers of accessions (more than 2000 accessions), the existing collections are analysed and described separately below (Sections 2.4.4, 2.4.5, 2.4.6, 2.4.7, 2.4.8, 2.4.9, 2.4.10). The remaining genera are treated together (Section 2.4.11), followed by an overview of *Brachypodium* genetic resources collections (Section 2.4.12).

2.4.3 Collections of Genetic Stocks and Mutants

Genetic stocks and mutants are plant genetic resources, along with modern and obsolete cultivars, landraces, wild relatives and breeding lines. They are gaining importance as tools in sophisticated applications of modern biotechnology in cereal improvement (CIMMYT 2007). Surveys of genetic stocks collections of wheat and barley were conducted in the frame of the respective “Crop Strategies” of the Global Crop Diversity Trust (CIMMYT 2007; Valkoun 2008), with the aim of ascertaining their size and status, and understanding how they could be best included in a global long-term conservation system. Summary information on genetic stocks and mutant collections can also be found in the two main databases consulted for this survey (Bioversity 2007; FAO 2007), but often these are listed together with other types of germplasm, and consequently, good estimates of their sizes are difficult to obtain. Surveys of European genetic stocks collections of wheat and barley are planned by the ECPGR Cereals Network for 2009–2010.

Genebanks usually give little emphasis on maintaining genetic stocks collections, since (1) they were regarded principally as research and teaching tools of low interest to plant breeding, (2) they require specialised inputs and

Table 2.3 Twenty-seven genebanks with more than 10,000 accessions of *Triticaceae*. For explanation of the FAO institution codes, see Section 2.2.2. The column Total gives the totals resulting from all sources included in this study. Abbreviations cf. Table 2.2

FAO code, Institute	Total	Trit.	Hord.	×Tr.	Aeg.	Sec.	Ely.	Agr.	Oth.
USA029, USDA-ARS, National Small Grains Collection, Aberdeen, Idaho, USA	94973	57053	30611	2023	2209	2114	0	0	963
MEX002, CIMMYT, El Batán, Mexico	94576	77462	11202	5427	6	0	0	0	479
SYR002, ICARDA, Aleppo, Syria	67285	36798	26018	6	4265	30	0	3	165
RUS001, Vavilov Institute of Plant Industry, St. Petersburg, Russia	64283	34808	20393	3744	2407	2914	0	0	17
DEU146, Leibniz Institute of Plant Genetics and Crop Plant Research, Gatersleben, Germany	54725	26833	22097	1571	1526	2392	123	25	158
CHN001, Institute of Crop Germplasm Resources, Chinese Academy of Agricultural Sciences, Beijing, China	54402	36794	17045	3	185	0	128	155	92
CAN004, Plant Gene Resources of Canada, Saskatoon, Saskatchewan, Canada	52952	10223	39928	113	874	1440	66	66	242
AUS003, Australian Winter Cereals Collection, Tamworth, Australia	37934	23811	13032	0	696	106	0	287	2
GBR011, John Innes Centre, Norwich, UK	35266	11058	23766	0	315	105	0	0	22
BRA003, EMBRAPA CENARGEN, Brasília, Brazil	34932	5131	29217	461	96	26	0	0	1
ITA004, Istituto del Germoplasma, Bari, Italy	30824	27744	1299	0	1401	214	0	0	166
ETH001, Institute of Biodiversity Conservation (IBC), Addis Ababa, Ethiopia	28904	13454	15450	0	0	0	0	0	0
UKR044, Institute of Agroecology and Biotechnology, Kyiv, Ukraine	28000	20000	8000	0	0	0	0	0	0
POL003, Plant Breeding and Acclimatization Institute, Radzikow, Poland	22135	11359	6088	2118	314	2226	8	3	19
BGR001, Institute for Plant Genetic Resources, Sadovo, Bulgaria	21303	12382	6003	926	631	1222	27	68	44
IRN029, National Plant Gene Bank of Iran, Karaj	17894	12169	5006	0	482	204	10	23	0
HUN003, Institute for Agrobotany, Tápíószele, Hungary	17648	8445	8297	257	162	335	41	87	24

Table 2. 3 (continued)

FAO code, Institute	Total	Trit.	Hord.	×Tr.	Aeg.	Sec.	Ely.	Agr.	Oth.
IND001, National Bureau of Plant Genetic Resources (NBPGR), New Delhi	17078	16440	500	125	0	0	0	13	0
IND069, All Indian Coordinated Wheat Programme, New Delhi, India	17000	17000	0	0	0	0	0	0	0
CZE122, Crop Research Institute, Prague, Czech Republic	16784	10424	4553	491	971	78	79	41	147
ISR003, Lieberman Germplasm Bank, Tel-Aviv University, Israel	16500	5500	8500	0	2500	0	0	0	0
BRA015, Centro Nacional de Pesquisa de Trigo, EMBRAPA, Passo Fundo, Brazil	16429	13133	2267	262	533	106	25	10	93
FRA040, Station d'Amélioration des Plantes, INRA, Clermont-Ferrand, France	14404	10715	3399	240	0	50	0	0	0
UKR001, National Centre for Plant Genetic Resources of Ukraine, Kharkiv, Ukraine	13703	7219	4394	1748	0	342	0	0	0
JPN003, National Institute of Agrobiological Science, Tsukuba, Japan	13455	7148	6242	34	0	31	0	0	0
CYP004, National Genebank, Nicosia, Cyprus	11329	7696	3433	0	200	0	0	0	0
CAN091, Soil and Crops Research and Development Centre, Sainte-Foy, Québec, Canada	10700	3700	5000	2000	0	0	0	0	0

management strategies, and (3) the development of genetic stocks was considered expensive and time-consuming and depending on strong interaction with sophisticated research programmes (CIMMYT 2007). Consequently, genetic stock collections are generally regarded as the responsibility of the individual or programme that initially created and maintained them. Such collections, however, have grown in size and importance, and efforts to better document and conserve them have been started. The Wheat Strategy (CIMMYT 2007) concludes that the role of genetic stocks in the global conservation effort of genetic resources should be re-evaluated, and they should be afforded a higher priority in the future. This is supported by the results of a survey among users of genebank collections about the most serious deficiencies in the current conservation of wheat genetic resources (CIMMYT 2007) that revealed that they miss sufficient access to mapping populations, isogens and genetic stocks; this is probably similar for the other cereals.

Some information about existing genetic stocks collections is provided below under the respective genera.

2.4.4 *Triticum*

For wheat, the combined databases included in the present analysis comprise a total of 727,000 accessions worldwide, of these, 167,000 accessions are named only to genus. Several *Triticum* species are cultivated, and there are many wild species of *Triticeae* that can be considered wild relatives.

The largest genebanks are CIMMYT with 10.6% of the global holdings, the USDA Small Grains Collection with 7.8%, and the Chinese Institute of Crop Germplasm Resources and ICARDA with 5.0% each, respectively (Table 2.4). The *Triticum* accessions belong to 35 species, including four hybrid species (Table 2.5). Global base collections are maintained by CIMMYT (for *T. aestivum*) and ICARDA (for *T. durum* and wheat wild relatives) (FAO 1996, p. 413).

The Global Crop Diversity Trust's "Crop Strategies" for wheat and barley contain overviews of germplasm collections worldwide, based on the FAO WIEWS database (FAO 2007), expert consultations and recent questionnaires. The Wheat Strategy (CIMMYT 2007) was developed in 2005–2007 under the coordination of D. Marshall and T. Payne and includes also triticale, *Aegilops* and rye. Since this strategy gives numbers of accessions per genebank mostly on the genus level only, the corresponding data are given here for comparison (Table 2.4), but could not be included in the present study.

Databases. No global wheat genetic resources database exists at present. CIMMYT is compiling the International Wheat Information System (IWIS) for passport data (FAO 1996). The CGIAR centres are working towards developing "Global Crop Registers" for several major crops that will cover also major non-CGIAR genebanks. A global database on genetic resources of

Table 2.4 Twenty-one institutions with more than 10,000 accessions of *Triticum*. The last column gives, for comparison, the figures from the Wheat Strategy of the Global Crop Diversity Trust (CIMMYT 2007). The last institute is included because it has more than 10,000 accessions according to the Wheat Strategy

FAO code	Institute	<i>Triticum</i> accs.	Wheat Strategy
MEX002	CIMMYT, El Batán, Mexico	77462	111681
USA029	USDA-ARS, National Small Grains Collection, Aberdeen, Idaho, USA	57053	56218
SYR002	ICARDA, Aleppo, Syria	36798	37830
CHN001	Institute of Crop Germplasm Resources, Chinese Academy of Agricultural Sciences, Beijing, China	36794	9633
RUS001	Vavilov Institute of Plant Industry, St. Petersburg, Russia	34808	39880
ITA004	Istituto di Genetica Vegetale, Bari, Italy	27744	32751
DEU146	Leibniz Institute of Plant Genetics and Crop Plant Research, Gatersleben, Germany	26833	9633
AUS003	Australian Winter Cereals Collection, Tamworth, Australia	23811	23917
UKR044	Institute of Agroecology and Biotechnology, Kyiv, Ukraine	20000	9620
IND069	All Indian Coordinated Wheat Programme, New Delhi, India	17000	17000
IND001	National Bureau of Plant Genetic Resources, New Delhi, India	16440	32880
ETH001	Institute of Biodiversity Conservation (IBC), Addis Ababa, Ethiopia	13454	10745
BRA015	Centro Nacional de Pesquisa de Trigo, EMBRAPA, Passo Fundo, Brazil	13133	13594
BGR001	Institute for Plant Genetic Resources, Sadovo, Bulgaria	12382	9747
IRN029	National Plant Gene Bank of Iran, Karaj	12169	12169
POL003	Plant Breeding and Acclimatization Institute, Radzików, Poland	11359	12974
GBR011	John Innes Centre, Norwich, UK	11058	9584
FRA040	Station d'Amélioration des Plantes, INRA, Clermont-Ferrand, France	10715	14200
CZE122	Crop Research Institute, Prague, Czech Republic	10424	11018
CAN004	Plant Gene Resources of Canada, Saskatoon, Saskatchewan, Canada	10223	5052
UKR001	National Centre for Plant Genetic Resources of Ukraine, Kharkiv, Ukraine	7219	20626

wild wheat and *Aegilops* species in genebanks was compiled since the end of the 1980s (Hodgkin et al. 1992; Konopka and Valkoun 2005). It is being maintained by ICARDA, contains data on almost 19,000 accessions from 62 genebanks, and is available on CD-ROM (Konopka, pers. comm. 2008).

In the framework of the European Cooperative Programme for Plant Genetic Resources (ECPGR), a Wheat Working Group forming part of the

Table 2.5 Total number of accessions of individual *Triticum* species, as reported in the combined database used for the present analysis. *Triticum* hybrid is a collective designation for named and unnamed hybrids between *Triticum* species. Additional 167,133 accessions are named only to genus (cf. Table 2.1)

Taxon	Accs.	Genebanks
<i>Triticum aestivum</i>	437744	152
<i>T. aethiopicum</i>	909	17
<i>T. araraticum</i>	1294	24
<i>T. boeoticum</i>	6087	38
<i>T. carthlicum</i>	672	38
<i>T. dicoccoides</i>	8028	42
<i>T. dicoccon</i>	4775	52
<i>T. durum</i>	49653	82
<i>T. erebuni</i>	5	2
<i>T. × flaksbergeri</i>	1	1
<i>T. × fungicidum</i>	26	9
<i>Triticum</i> hybrid	367	4
<i>T. ispahanicum</i>	53	16
<i>T. jakubzineri</i>	5	5
<i>T. karamyshevii</i>	71	25
<i>T. kiharae</i>	8	6
<i>T. macha</i>	232	28
<i>T. × miguschovae</i>	5	3
<i>T. militinae</i>	17	12
<i>T. monococcum</i>	5367	54
<i>T. palaeocolchicum</i>	4	1
<i>T. palmovae</i>	6	3
<i>T. petropavlovskiyi</i>	16	9
<i>T. polonicum</i>	573	39
<i>T. sinskajae</i>	37	12
<i>T. × soveticum</i>	1	1
<i>T. spelta</i>	5839	51
<i>T. sphaerococcum</i>	650	29
<i>T. × timococcum</i>	8	7
<i>T. timonovum</i>	2	1
<i>T. timopheevii</i>	590	37
<i>T. turanicum</i>	236	21
<i>T. turgidum</i>	39513	49
<i>T. urartu</i>	2186	20
<i>T. vavilovii</i>	84	25
<i>T. zhukovskiyi</i>	64	22

Cereals Network was established in 1988 (Lipman et al. 2005). The European Wheat Database (EWDB 2008) was developed by the Crop Research Institute in Prague, Czech Republic, covering 59 institutions in 32 countries with a total of 155,950 accessions of *Triticum* and 6,910 accessions of *Aegilops* (I. Faberová, pers. comm. Feb. 2008). An online database of genealogies and identified alleles

of genes of 74,527 wheat genotypes (Martynov et al. 2006a), based on 3,261 single information sources, is also accessible from Prague. Earlier publications of wheat pedigrees include Zeven and Zeven-Hissink (1976) and Martynov et al. (1992), with information on 14,000 and 31,000 cultivars, respectively.

Core Collections. For wheat, several core collections are reported in the literature. These include, among others:

- a core collection of 645 spring wheat accessions selected from 2,123 accessions kept in the Czech genebank (Stehno et al. 2006), as well as a winter wheat core collection (Martynov et al. 2003);
- a durum wheat core collection (Spagnoletti Zeuli and Qualset 1993, 1995);
- a core collection of wheat from the Serbian genebank in Novi Sad (Kobiljski et al. 2002);
- a *T. aestivum* core collection from Northwestern China (Hao et al. 2006);
- a *T. monococcum* core collection for eyespot resistance evaluation (Cadle et al. 1997).

Genetic Stocks Collections. A survey of genetic stocks collections in wheat was carried out in the frame of the Wheat Strategy (CIMMYT 2007). Eighteen collections from 11 countries are reported there, but several significant repositories in Japan, Russia, the USA (3), and the United Kingdom, as well as some key laboratories elsewhere, did not respond. Organisations such as the International Wheat Genetics Symposium, or the European Cereals Aneuploid Co-operative (EWAC, formerly European Cereals Aneuploid Co-operative), reveal important stocks and collections. Börner and Worland (1995) compiled an overview of 35 cytogenetic tester stocks collections from 19 countries, but without indicating their sizes. Taken together, the surveys of Börner and Worland (1995), CIMMYT (2007) and information from the combined FAO and Bioversity databases about collections of wheat genetic stocks and mutants (search categories GS and MT) reveal ca. 55 different institutions in 30 countries. For 28 collections with known number of accessions, the accumulated amount to 30,716 accessions (Table 2.6). The actual total size of wheat genetic stocks collections worldwide must be assumed even higher.

2.4.5 *Hordeum*

FAO (1996) counts 485,000 barley accessions worldwide. Hintum and Menting (2003) adjusted this figure to 370,796. The present study gives a total of 453,602 accessions. Most of these belong to the cultivated barley, *Hordeum vulgare* subsp. *vulgare*, and its close wild relative, subsp. *spontaneum*, forming the primary gene pool.

The largest barley collections are: PGRC Canada (8.8% of the global holdings), USDA (6.7%), EMBRAPA Brazil (6.4%) and ICARDA (5.7%). Genebanks with more than 8,000 accessions are given in Table 2.7. Totals for

Table 2.6 Twelve selected institutions holding large collections of wheat genetic stocks with more than 600 accessions

FAO code	Institute	Genetic Stocks	Source
USA109	Department of Botany and Plant Sciences, Riverside, California, USA	6785	Wheat Strategy
ITA	Milano, Italy	5985	Wheat Strategy
USA	USDA/ARS, Wheat Genetic Stocks Collection, University of Missouri, Columbia, Missouri	3000	Wheat Strategy
USA	Wheat Genetic and Genomic Resources Center (WGGRC), Kansas State University, Manhattan, Kansas, USA	2780	C. Feuillet, pers. comm. August 2008; WGGRC website
AUS003	Australian Winter Cereals Collection, Tamworth, Australia	2627	Wheat Strategy
USA320	North Dakota State University of Agriculture and Applied Science, Fargo, North Dakota, USA	1989	Wheat Strategy
CAN015	Cereal Research Centre, Agriculture and Agri-Food Canada, Winnipeg, Manitoba, Canada	1560	FAO, Bioversity
ALB001	Department of Agronomy, Agricultural University, Tirana, Albania	1123	FAO, Bioversity
HUN020	Agricultural Research Institute, Hungarian Academy of Sciences, Martonvásár, Hungary	651	Wheat Strategy, FAO, Bioversity, Kovács et al. 2002
GRC001	Cereal Institute, National Agricultural Research Foundation, Thessaloniki, Greece	645	FAO, Bioversity
CHE001	Agroscope Changins-Wädenswil (ACW), Switzerland	629	Wheat Strategy
FRA094	Station de Génétique Végétale, INRA, Gif-sur-Yvette, France	600	FAO, Bioversity
Total		30,716	

genebank collections as collated in the Barley Strategy of the Global Crop Diversity Trust (Valkoun 2008, unpublished, cf. above under Section 2.4.3 for wheat) are provided for comparison with our results in Table 2.7. The *Hordeum* accessions belong to 32 species (Table 2.8).

Hintum and Menting (2003) provided estimates of the degree of duplication between three of the largest barley collections (CAN004, USA029, and SYR002) and found that out of a total of >95,000 accessions, only ca. 51,000 seem to be different. When considering three medium-size collections (IPK and the former BAZ collections in Germany, the latter being recently merged into the IPK genebank, and CGN, The Netherlands, with 12,866, 7,604, and 3,485 reported accessions, respectively), they found that ca. 85% of these accessions were kept in only one of them, i.e. that these were relatively distinct (Hintum and Menting 2000, 2003).

Table 2.7 Twenty institutions with more than 8,000 accessions of *Hordeum*. The last column gives, for comparison, the figures from the Barley Strategy of the Global Crop Diversity Trust (Valkoun 2008). The last five institutions listed have more than 8,000 accessions in the Barley Strategy, but less in our survey

FAO code	Institute	<i>Hordeum</i> accs.	Barley Strategy
CAN004	Plant Gene Resources of Canada, Saskatoon, Saskatchewan, Canada	39928	39852
USA029	USDA-ARS, National Small Grains Collection, Aberdeen, Idaho, USA	30611	29838
BRA003	EMBRAPA CENARGEN, Brasília, Brazil	29217	29227
SYR002	ICARDA, Aleppo, Syria	26018	26117
GBR011	John Innes Centre, Norwich, UK	23766	23603
DEU146	Leibniz Institute of Plant Genetics and Crop Plant Research, Gatersleben, Germany	22097	22106
RUS001	Vavilov Institute of Plant Industry, St. Petersburg, Russia	20393	17850
CHN001	Institute of Crop Germplasm Resources, Chinese Academy of Agricultural Sciences, Beijing, China	17045	18818
ETH001	Institute of Biodiversity Conservation (IBC), Addis Ababa, Ethiopia	15450	15360
AUS003	Australian Winter Cereals Collection, Tamworth, Australia	13032	12600
MEX002	CIMMYT, El Batán, Mexico	11202	11202
AUS091	University of Western Australia, Crawley, Perth, W Australia	9031	
ISR003	Lieberman Germplasm Bank, Tel-Aviv University, Israel	8500	6662
HUN003	Institute for Agrobotany, Tápiószéle, Hungary	8297	4208
UKR044	Institute of Agroecology and Biotechnology, Kyiv, Ukraine	8000	
KOR003	National Institute of Agricultural Biotechnology, Suwon, Republic of Korea	7078	18764
JPN009	Research Institute for Bioresources, Okayama University, Kurashiki, Japan	5435	14106
SWE002	NordGen, Alnarp, Sweden	6333	13435
JPN003	National Institute of Agrobiological Science, Tsukuba, Japan	6242	8806
IND001	National Bureau of Plant Genetic Resources (NBPGR), New Delhi, India	500	8384

Databases. A “Global Barley Register” is being developed at ICARDA and will include the collections of CGIAR and important non-CGIAR genebanks (J. Konopka, pers. comm. 2008). In Europe, the ECPGR Barley Working Group has been one of the first crop working groups established in the early 1980s (reports can be found via the website of the ECPGR Cereals Network, cf. Appendix).

Table 2.8 Total number of accessions per *Hordeum* taxon, as reported in the combined database used for the present analysis. *Hordeum* hybrid is a collective designation for named and unnamed hybrids between *Hordeum* species. Additional 63,511 accessions are named only to genus (cf. Table 2.1)

Taxon	Accs.	Genebanks
<i>Hordeum arizonicum</i>	18	6
<i>H. bogdanii</i>	76	9
<i>H. brachyantherum</i>	78	6
subsp. <i>brachyantherum</i>	58	4
subsp. <i>californicum</i>	46	5
<i>H. brevisubulatum</i>	47	8
subsp. <i>brevisubulatum</i>	2	1
subsp. <i>iranicum</i>	6	1
subsp. <i>nevskianum</i>	1	1
subsp. <i>turkestanicum</i>	7	4
subsp. <i>violaceum</i>	63	11
<i>H. bulbosum</i>	1124	30
<i>H. capense</i>	14	8
<i>H. chilense</i>	274	13
<i>H. comosum</i>	63	10
<i>H. cordobense</i>	20	9
<i>H. depressum</i>	54	8
<i>H. erectifolium</i>	5	4
<i>H. euclaston</i>	32	7
<i>H. flexuosum</i>	28	6
<i>H. fuegianum</i>	20	4
<i>H. guatemalense</i>	1	1
<i>H. halophilum</i>	2	1
<i>Hordeum</i> hybrid	46	3
<i>H. intercedens</i>	33	6
<i>H. jubatum</i>	228	17
<i>H. lechleri</i>	111	9
<i>H. marinum</i>	97	12
subsp. <i>gussoneanum</i>	192	16
subsp. <i>marinum</i>	70	8
<i>H. murinum</i>	237	18
subsp. <i>glaucum</i>	184	12
subsp. <i>leporinum</i>	642	18
subsp. <i>murinum</i>	25	6
<i>H. muticum</i>	54	9
<i>H. parodii</i>	117	8
<i>H. patagonicum</i>	55	3
subsp. <i>magellanicum</i>	33	4
subsp. <i>mustersii</i>	4	3
subsp. <i>patagonicum</i>	12	3
subsp. <i>santacruzense</i>	27	4
subsp. <i>setifolium</i>	22	4
<i>H. procerum</i>	32	8
<i>H. pubiflorum</i>	107	6

Table 2.8 (continued)

Taxon	Accs.	Genebanks
<i>H. pusillum</i>	58	10
<i>H. roshevitzii</i>	27	8
<i>H. secalinum</i>	108	13
<i>H. stenostachys</i>	110	12
<i>H. tetraploidum</i>	11	4
<i>H. vulgare</i>	42251	43
subsp. <i>spontaneum</i>	32385	37
subsp. <i>vulgare</i>	299165	133
var. <i>agriocrithon</i>	1690	16
var. <i>lagunculiforme</i>	9925	5

The European Barley Database (EBDB) was established by the Leibniz Institute of Plant Genetics and Crop Plant Research in Gatersleben, Germany, since 1983 (e.g., Knüpfper 1988; Enneking and Knüpfper 2001). It contains passport data of 155,558 accessions of 37 barley species and more than 200 botanical varieties, maintained in 43 European and Non-European institutions, and includes the International Barley Core Collection (BCC, see below).

Pedigrees of 6,443 barley cultivars and lines, based on 99 information sources, can be found online on the server of the Prague genebank (Martynov et al. 2006b). The most up-to-date version contains 7579 records from 136 sources but is not online yet (S. Martynov, pers. comm. 2008). Published registers of barley cultivars include Arias et al. (1983), Baum et al. (1985), and Baumer and Cais (2000).

Core Collections. The International Barley Core Collection (BCC) developed and maintained at several institutions worldwide consists of ca. 1,500 accessions covering *H. vulgare* subsp. *vulgare* (landraces and cultivars), subsp. *spontaneum* and the other wild species as well as genetic stocks (Knüpfper and Hintum 1995, 2003). With respect to genetic stocks, only material which is not too difficult to regenerate by genebank staff was included in the BCC.

Other persistent barley core collections (i.e. which continue to exist after their establishment) reported in the literature include:

- the Barley Core Subset of the USDA collection developed at the USDA-ARS National Small Grain Collection (NSGC) to assist scientist studying diversity in barley and focussing new evaluations (Bockelman 1996). It consists of 2,303 accessions (H. Bockelman, pers. comm. 2002);
- the ICARDA Barley Core Collection consisting of 153 *Hordeum vulgare* accessions selected from the ICARDA barley holdings (Chabane and Valkoun 2004);
- the Spanish Barley Core Collection, consisting of ca. 160 landraces collected in Spain prior to the introduction of modern cultivars selected from ca. 2,000 accessions (Igartua et al. 1998). This collection is being intensively studied (e.g. Yahiaoui et al. 2008);
- the Uruguayan Barley Core Collection of malting types consisting of ca. 45 accessions (Malosetti et al. 2000).

Genetic Stocks Collections. A survey of barley genetic stocks collections was carried out in the frame of the Barley Strategy (Valkoun 2008). Seventeen collections from 17 countries totalling 25,300 accessions of genetic stocks are reported there. Of ca. 291,000 accessions with known germplasm type, 9% are genetic stocks. According to this survey, the largest collection is held at NordGen in Alnarp, Sweden, with ca. 10,000 accessions that resulted from a joint Scandinavian mutation research programme, and 685 translocation and 58 duplication lines. Other large genetic stocks collections are maintained at PGRC Canada, USDA, and NIAR Japan.

At NordGen (Sweden), a Database on Barley Genes and Genetic Stocks (BGS) (Lundqvist and Huldén 2005) was developed using AceDB, the database software used also for GrainGenes. BGS will be included in GrainGenes and provides the possibility of including more data on genetic stocks collections (Valkoun 2008).

Organisations such as the International Barley Genetics Symposium reveal important stocks and collections (e.g., Lundqvist 2007 and a series of earlier annual reports in Barley Genetics Newsletter where information about many genetic stocks collections can be found). A recent survey of European barley collections, includes also genetic stocks and mutant collections (Knüpffer 2008, unpublished). Taken together, the surveys of Valkoun (2008), Knüpffer (2008) and information from the combined FAO and Bioversity databases about collections of barley genetic stocks and mutants (search categories GS and MT) provide information on ca. 40 different institutions in 26 countries. For 29 collections with known number of accessions, the accumulated sizes amount to 33,671 accessions (Table 2.9). The actual total size of barley genetic stocks collections worldwide must be assumed even higher.

2.4.6 × *Triticosecale*

Within the hybrid genus ×*Triticosecale*, species names are not applied. Of the 31,867 accessions in our study, almost two-thirds are held in European genebanks, but the largest single collection is held by CIMMYT (17% of the world holdings), followed by the Vavilov Institute in St. Petersburg, Russia (11.7%), the Polish genebank (6.6%) and USDA (6.3%) (Table 2.10). Triticale is also reflected in the Wheat Strategy of the Global Crop Diversity Trust (CIMMYT, 2007), which estimates about 35,000 accessions globally, of which 10 major genebanks account to a total of 32,000 accessions.

A network on triticale genetic resources was created (Darvey et al. 1996). Within ECPGR, a meeting on rye and triticale genetic resources was held recently (Kleijer et al. 2007), but a formal Working Group does not exist.

Databases. The European Triticale Database (ETDB) was established at Agro-scope Changins-Wädenswil (ACW), Nyon, Switzerland, following a recommendation of the ECPGR Wheat Genetic Resources Workshop in 1996. Data have

Table 2.9 Ten selected institutions holding large collections of barley genetic stocks with more than 500 accessions

FAO code	Institute	Genetic stocks	Source
SWE002	NordGen, Alnarp, Sweden	10024	Barley Strategy, ECPGR
CAN004	Plant Gene Resources of Canada, Saskatoon, Saskatchewan, Canada	6011	Barley Strategy
USA005	Barley Genetic Stocks Center, National Seed Storage Laboratory, USDA-ARS, Fort Collins, Colorado, USA	3203	Barley Strategy
CAN005	Département de phytologie, Université Laval, Québec, Québec, Canada	3000	FAO, Bioversity
JPN009	Research Institute for Bioresources, Okayama University, Kurashiki, Japan	2890	Barley Strategy
DEU146	Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Gatersleben, Germany	2562	Barley Strategy, FAO, Bioversity ¹
ITA024	Experimental Institute for Cereal Crops, Fiorenzuola d'Arda, Italy	1240	Barley Strategy, FAO, Bioversity, ECPGR
BRA015	Centro Nacional de Pesquisa de Trigo, EMBRAPA, Passo Fundo, Brazil	860	Barley Strategy
RUS001	Vavilov Institute of Plant Industry, St. Petersburg, Russia	579	Barley Strategy
SYR002	ICARDA, Aleppo, Syria	551	Barley Strategy
Total		33,671	

¹ Figure updated according to GBIS (2008); some additional IPK collections exist outside the genebank.

Table 2.10 Eighteen institutions with over 500 accessions of *× Triticosecale*. The last column gives, for comparison, the figures from the Wheat Strategy of the Global Crop Diversity Trust (CIMMYT 2007). Two collections are listed in the Wheat strategy as having above 500 accessions, whereas in our survey they have less.

FAO code	Institute	<i>× Triticosecale</i> accs.	Wheat Strategy
MEX002	CIMMYT, El Batán, Mexico	5427	17871
RUS001	Vavilov Institute of Plant Industry, St. Petersburg, Russia	3744	4593
POL003	Plant Breeding and Acclimatization Institute, Radzików, Poland	2118	1586
USA029	USDA-ARS, National Small Grains Collection, Aberdeen, Idaho, USA	2023	2007
CAN091	Soil and Crops Research and Development Centre, Agriculture and Agri-Food Canada Sainte-Foy, Quebec, Canada	2000	
POL025	Institute of Genetics and Plant Breeding, University of Agriculture, Lublin, Poland	1748	1748

Table 2.10 (continued)

FAO code	Institute	× <i>Triticosecale</i> accs.	Wheat Strategy
UKR001	National Centre for Plant Genetic Resources of Ukraine, Kharkiv, Ukraine	1748	2184
DEU146	Leibniz Institute of Plant Genetics and Crop Plant Research, Gatersleben, Germany	1571	502
BGR001	Institute for Plant Genetic Resources, Sadovo, Bulgaria	926	
SVK001	Research Institute of Plant Production Piestany, Slovak Republic	800	
CHE001	Agroscope Changins-Wädenswil (ACW), Switzerland	750	750
ROM002	Research Institute for Cereals and Technical Plants, Fundulea, Romania	629	
PRT076	Estação Nacional Melhoramento Plantas, Elvas, Portugal	600	
ESP004	Centro de Recursos Fitogenéticos, Alcalá de Henares, Spain	511	
AUT011	Austrian Agency for Health and Food Safety / Plant Production, Vienna, Austria	510	
ESP013	Universitat de Lleida e Institut de Recerca i Tecnologia Agroalimentàries, Centre UdL-IRTA, Lleida, Spain	501	
USA288	Department of Agronomy, University of Missouri, Columbia, USA		1400
FRA040	Plant Breeding and Genetics Department, INRA, Clermont-Ferrand, France	240	800

been collected from 22 institutes in 18 different countries (Kleijer 2007). These institutes conserve a total of 11,708 accessions (ETDB 2008). A triticale cultivar register database with information on pedigrees, synonyms, breeders and references was developed by Baum et al. (1990), covering ca. 1,100 cultivar names (Baum 1995), but it is not online accessible (B.R. Baum, pers. comm. 2008).

Genetic Stocks Collections. Six collections of triticale genetic stocks in Germany, Greece, Hungary and Romania are documented by the databases of FAO and Bioversity, totalling 450 accessions, the three larger ones being presented in Table 2.11.

2.4.7 *Aegilops*

Germplasm collections of the genus *Aegilops* are often dealt with together with *Triticum*. For example, the Wheat Strategy (CIMMYT, 2007) includes *Aegilops*, and its collections are also included in the European Wheat Database of the ECPGR and the Wild Wheat and *Aegilops* Database (Konopka and Valkoun 2005). However, detailed information about *Aegilops* collections is not included in the Wheat Strategy.

Table 2.11 Three institutions with more than 50 accessions of triticale genetic stocks, according to the FAO and Bioversity databases

FAO code	Institute	Genetic stocks
HUN017	Faculty of Agricultural Sciences, Pannon University of Agriculture, Mosonmagyaróvár, Hungary	229
DEU004	State Plant Breeding Institute, University of Hohenheim, Stuttgart, Germany	150
GRC001	Cereal Institute, National Agricultural Research Foundation, Themi-Thessaloniki, Greece	57
Total		450

In the present study, 31,401 accessions of *Aegilops* from 61 genebanks are documented. The largest collections are held by ICARDA (with 13.6% of the global holdings), Tel Aviv University, Israel (8.0%), the Vavilov Institute, St. Petersburg, Russia (7.7%), and the Plant Germplasm Institute, Kyoto University, Japan (7.6%) (Table 2.12). *Amblyopyrum muticum* (138 accessions),

Table 2.12 Nineteen institutions with over 500 accessions of *Aegilops*

FAO code	Institute	<i>Aegilops</i> accs.
SYR002	ICARDA, Aleppo, Syria	4265
ISR003	Lieberman Germplasm Bank, Tel-Aviv University, Israel	2500
RUS001	Vavilov Institute of Plant Industry, St. Petersburg, Russia	2407
JPN001	Plant Germplasm Institute, Kyoto University, Japan	2396
USA029	USDA-ARS, National Small Grains Collection, Aberdeen, Idaho, USA	2209
DEU146	Leibniz Institute of Plant Genetics and Crop Plant Research, Gatersleben, Germany	1526
USA	Wheat Genetic and Genomic Resources Center (WGGRC), Kansas State University, Manhattan, Kansas, USA	1425
ITA004	Istituto di Genetica Vegetale, Bari, Italy	1401
CAN015	Cereal Research Centre, Agriculture and Agri-Food Canada, Winnipeg, Manitoba, Canada	1100
FRA010	Station d'Amélioration des Plantes, INRA, Le Rheu, France	1070
GRC005	Greek Genebank, Themi, Thessaloniki, Greece	978
CZE122	Crop Research Institute, Prague, Czech Republic	971
ARM002	Armenian State Agrarian University, Laboratory of Plant Gene Pool and Breeding, Yerevan, Armenia	941
CAN004	Plant Gene Resources of Canada, Saskatoon, Saskatchewan, Canada	874
USA109	Department of Botany and Plant Sciences, Riverside, California, USA	769
SYR055	Genetic Resources Department, General Commission for Scientific Agricultural Research, Damascus, Syria	743
AUS003	Australian Winter Cereals Collection, Tamworth, Australia	696
BGR001	Institute for Plant Genetic Resources, Sadovo, Bulgaria	631
BRA015	Centro Nacional de Pesquisa de Trigo, EMBRAPA, Passo Fundo, Brazil	533

Table 2.13 Total number of accessions per *Aegilops* species, as reported in the combined database used for the present analysis. *Aegilops* hybrid is a collective designation for named and unnamed hybrids between *Aegilops* species. Additional 10,340 accessions are named only to genus (cf. Table 2.1)

Taxon	Accs.	Genebanks
<i>Aegilops bicornis</i>	260	21
<i>Ae. biuncialis</i>	995	14
<i>Ae. columnaris</i>	292	25
<i>Ae. comosa</i>	281	17
<i>Ae. crassa</i>	700	26
<i>Ae. cylindrica</i>	2547	34
<i>Ae. geniculata</i>	1793	30
<i>Aegilops</i> hybrid	26	3
<i>Ae. juvenalis</i>	104	20
<i>Ae. kotschy</i>	219	21
<i>Ae. ligustica</i>	28	4
<i>Ae. longissima</i>	279	23
<i>Ae. lorentii</i>	1244	19
<i>Ae. markgrafii</i>	460	20
<i>Ae. neglecta</i>	1145	30
<i>Ae. peregrina</i>	410	17
<i>Ae. recta</i>	21	2
<i>Ae. searsii</i>	201	13
<i>Ae. sharonensis</i>	124	11
<i>Ae. speltoides</i>	1125	32
<i>Ae. tauschii</i>	3828	37
<i>Ae. triaristata</i>	74	8
<i>Ae. triuncialis</i>	3833	34
<i>Ae. umbellulata</i>	434	24
<i>Ae. uniaristata</i>	91	12
<i>Ae. variabilis</i>	8	4
<i>Ae. vavilovii</i>	182	8
<i>Ae. ventricosa</i>	346	28

sometimes included in *Aegilops*, is treated under “other *Triticeae*” below. The *Aegilops* accessions documented belong to 27 species (Table 2.13).

Databases. A global database on genetic resources of wild wheat and *Aegilops* species in genebanks is mentioned above under *Triticum*. Using this database together with three others, Maxted et al. (2008) compiled and evaluated a database of 9,866 unique georeferenced *Aegilops* accessions from four major databases on genetic resources, with the aim to reveal possible gaps in *ex situ* collections and to propose *in situ* conservation measures. The European Wheat Database of the ECPGR reports 6,910 accessions of *Aegilops* (cf. Section 2.4.4, *Triticum*).

Core Collection. Zaharieva et al. (2001) describe the establishment of a core collection of *Ae. geniculata*.

2.4.8 *Secale*

The Wheat Strategy of the Global Crop Diversity Trust (CIMMYT 2007) reports that there are more than 20,000 rye accessions in genebanks, of which ca. 10,000 accessions are held in Europe. Our study documents 21,550 accessions worldwide, with 15,845 in Europe. The largest collections are the Vavilov Institute in St. Petersburg, Russia (13.5%), the Gatersleben Genebank in Germany (11.1%), the Polish Genebank in Radzików (10.3%) and the USDA Small Grains Collection (9.8%) (Table 2.14). The internal duplication in the Gatersleben collection is probably rather high, since it was brought together from three former German collections for which the overlap was estimated by

Table 2.14 Seventeen institutions with more than 300 accessions of *Secale*. The last column gives, for comparison, the figures from the Wheat Strategy of the Global Crop Diversity Trust (CIMMYT 2007). The rye collection of CIMMYT is not documented in the information sources used in our study but has more than 300 accessions according to the Wheat Strategy

FAO code	Institute	<i>Secale</i> accs.	Wheat Strategy
RUS001	Vavilov Institute of Plant Industry, St. Petersburg, Russia	2914	2685
DEU146	Leibniz Institute of Plant Genetics and Crop Plant Research, Gatersleben, Germany	2392	2154
POL003	Plant Breeding and Acclimatization Institute, Radzików, Poland	2226	1354
USA029	USDA-ARS, National Small Grains Collection, Aberdeen, Idaho, USA	2114	1897
POL001	Botanical Garden, Polish Academy of Sciences, Warsaw, Poland	1631	1630
CAN004	Plant Gene Resources of Canada, Saskatoon, Saskatchewan, Canada	1440	1440
BGR001	Institute for Plant Genetic Resources, Sadovo, Bulgaria	1222	400
TUR001	Plant Genetic Resources Department, Aegean Agricultural Research Institute, Izmir, Turkey	851	585
CZE047	Agricultural Research Institute Kroměříž, Ltd., Kroměříž, Czech Republic	695	663
ESP004	Centro de Recursos Fitogenéticos, Alcalá de Henares, Spain	487	412
SWE002	NordGen, Alnarp, Sweden	382	365
PRT005	Banco de Germoplasma, Instituto Nacional de Investigação Agrária, Oeiras, Portugal	377	580
ROM028	Agricultural Research Station Suceava, Romania	343	
UKR001	National Centre for Plant Genetic Resources of Ukraine, Kharkiv, Ukraine	342	
HUN003	Institute for Agrobotany, Tápiószéle, Hungary	335	361
DEU013	Institute of Plant Breeding, University of Hohenheim, Stuttgart, Germany	300	
MEX002	CIMMYT, El Batán, Mexico		747

Table 2.15 Total number of accessions per *Secale* species, as reported in the combined database used for the present analysis. *Secale* hybrid is a collective designation for named and unnamed hybrids between *Secale* species. Additional 2,069 accessions are named only to genus (cf. Table 2.1)

Taxon	Accs.	Genebanks
<i>Secale cereale</i> subsp. <i>afghanicum</i>	26	7
subsp. <i>ancestrale</i>	37	8
subsp. <i>cereale</i>	18410	78
subsp. <i>dalmaticum</i>	2	1
subsp. <i>dighoricum</i>	19	9
subsp. <i>rigidum</i>	3	3
subsp. <i>segetale</i>	92	13
subsp. <i>tetraploidum</i>	267	8
subsp. <i>tsitsinii</i>	3	3
var. <i>multicaule</i>	10	4
<i>S.</i> × <i>derzhavinii</i>	18	4
<i>Secale</i> hybrid	37	1
<i>S. segetale</i>	1	1
<i>S. strictum</i> subsp. <i>africanum</i>	22	10
subsp. <i>anatolicum</i>	38	9
subsp. <i>ciliatoglume</i>	4	4
subsp. <i>kuprijanovii</i>	63	11
subsp. <i>strictum</i>	217	25
<i>S. sylvestre</i>	129	17
<i>S. vavilovii</i>	83	16

Maschka (1995) when these were still separate. Six species of *Secale* are kept in genebanks (Table 2.15).

Databases. The European Rye Database of the ECPGR was established since 1982 in Poland, as the first of its kind (Podyma 1998). At present, it documents 13,610 accessions of 36 genebanks in 26 countries (Zaczyński 2007). Like with triticale, ECPGR does not have a formal working group on rye, but meetings of rye genetic resources specialists were organised in Europe (e.g., Gass et al. 1998; Kleijer et al. 2007).

Genetic Stocks Collections. The FAO and Bioversity databases indicate three rye genetic stocks collections in Germany, Greece and Hungary, with a total of 35 accessions (Table 2.16).

Table 2.16 Rye genetic stocks collections according to FAO and Bioversity databases

FAO code	Institute	Genetic stocks
DEU146	Leibniz Institute of Plant Genetics and Crop Plant Research, Gatersleben, Germany	20
GRC001	Cereal Institute, National Agricultural Research Foundation, Themi-Thessaloniki, Greece	15
HUN003	Institute for Agrobotany, Tápíószele, Hungary	

2.4.9 *Elymus*

The genus *Elymus* is represented in global genebanks by 4,132 accessions. A single collection, the Swedish Agricultural University in Alnarp, holds more than half of this number (53.2%). Following in size of their *Elymus* collections are the Western Regional Plant Introduction Station Pullman, USA (16.5%) and the Nordic Gene Bank in Alnarp, Sweden (6.9%) (Table 2.17). The *Elymus* accessions preserved in genebanks belong to 87 species (Table 2.18).

Table 2.17 Nine institutions with more than 50 accessions of *Elymus*

FAO code	Institute	<i>Elymus</i> accs.
SWE001	Department of Plant Breeding Research, Swedish University of Agricultural Sciences, Alnarp, Sweden	2198
USA022	Western Regional Plant Introduction Station, USDA-ARS, Washington State University, Pullman, Washington, USA	681
SWE002	NordGen, Alnarp, Sweden	287
USA169	Forage and Range Research USDA/ARS, Grass Collection, Logan, Utah, USA	130
CHN001	Institute of Crop Germplasm Resources, Chinese Academy of Agricultural Sciences, Beijing, China	128
DEU146	Leibniz Institute of Plant Genetics and Crop Plant Research, Gatersleben, Germany	123
CZE122	Crop Research Institute, Prague, Czech Republic	79
CAN004	Plant Gene Resources of Canada, Saskatoon, Saskatchewan, Canada	66
ESP004	Centro de Recursos Fitogenéticos, Alcalá de Henares, Spain	54

Table 2.18 Total number of accessions per *Elymus* taxon, as reported in the combined database used for the present analysis. Additional 2,585 accessions are named only to genus (cf. Table 2.1)

Taxon	Accs.	Genebanks
<i>Elymus abolinii</i>	14	3
<i>E. agropyroides</i>	8	2
<i>E. alashanicus</i>	1	1
<i>E. alaskanus</i>	11	1
<i>E. andinus</i>	9	2
<i>E. antarcticus</i>	31	3
<i>E. antiquus</i>	9	2
<i>E. apricus</i>	2	1
<i>E. araucanus</i>	1	1
<i>E. arizonicus</i>	2	2
<i>E. austromontanus</i>	7	1
<i>E. bakeri</i>	7	1
<i>E. barbicallus</i>	5	2
<i>E. borianus</i>	3	2

Table 2.18 (continued)

Taxon	Accs.	Genebanks
<i>E. brachyaristatus</i>	1	1
<i>E. breviaristatus</i>	17	3
<i>E. buschianus</i>	2	2
<i>E. canadensis</i>	107	7
<i>E. canaliculatus</i>	1	1
<i>E. caninus</i>	134	12
<i>E. caucasicus</i>	2	1
<i>E. ciliaris</i>	20	5
var. <i>amurensis</i>	2	1
<i>E. confusus</i>	9	3
<i>E. cordilleranus</i>	3	1
<i>E. curvatus</i>	2	1
<i>E. dahuricus</i>	373	13
<i>E. dentatus</i>	7	1
<i>E. donianus</i>	2	2
<i>E. drobovii</i>	10	4
<i>E. elymoides</i>	80	2
<i>E. enysii</i>	2	2
<i>E. excelsus</i>	4	3
<i>E. falcis</i>	1	1
<i>E. fedtschenkoi</i>	16	3
<i>E. fibrosus</i>	30	5
<i>E. foliosus</i>	2	2
<i>E. gayanus</i>	9	3
<i>E. glaucissimus</i>	7	1
<i>E. glaucus</i>	61	6
<i>E. gmelinii</i>	28	7
<i>E. gracilis</i>	1	1
<i>E. hitchcockii</i>	1	1
<i>E. hordeoides</i>	1	1
<i>E. hystrix</i>	4	4
<i>E. jacquemontii</i>	2	1
<i>E. jacutensis</i>	5	2
<i>E. komarovii</i>	1	1
<i>E. lanceolatus</i>	48	7
<i>E. longearistatus</i>	4	2
<i>E. macrochaetus</i>	1	1
<i>E. macrourus</i>	3	1
<i>E. magellanicus</i>	13	3
<i>E. mutabilis</i>	27	3
<i>E. nakaii</i>	2	1
<i>E. nevskii</i>	9	3
<i>E. nutans</i>	31	6
<i>E. oschensis</i>	1	1
<i>E. panormitanus</i>	6	4
<i>E. parodii</i>	1	1

Table 2.18 (continued)

Taxon	Accs.	Genebanks
<i>E. patagonicus</i>	12	7
<i>E. pendulinus</i>	10	1
subsp. <i>brachypodioides</i>	7	2
<i>E. praeeruptus</i>	3	1
<i>E. pseudonutans</i>	1	1
<i>E. purpurascens</i>	1	1
<i>E. repens</i>	48	15
<i>E. scaber</i>	6	2
<i>E. scabrifolius</i>	8	3
<i>E. scabriglumis</i>	4	1
<i>E. schrenkianus</i>	2	1
<i>E. scribneri</i>	3	3
<i>E. semicostatus</i>	11	4
<i>E. sibiricus</i>	103	11
<i>E. strictus</i>	11	2
<i>E. tianshanicus</i>	3	1
<i>E. tibeticus</i>	1	1
<i>E. trachycaulus</i>	41	14
<i>E. transbaicalensis</i>	11	1
<i>E. transhyrcanus</i>	3	3
<i>E. truncatus</i>	1	1
<i>E. tschimganicus</i>	7	4
<i>E. tsukushiensis</i>	5	4
<i>E. uralensis</i> subsp. <i>komarovii</i>	5	1
<i>E. vaillantianus</i>	2	2
<i>E. villosus</i>	2	2
<i>E. violaceus</i>	11	2
<i>E. virginicus</i>	9	3
<i>E. wiegandii</i>	2	2

2.4.10 *Agropyron*

The genus *Agropyron* has 3,110 accessions in genebanks worldwide. About half of this number is documented only to the level of genus. Since many species nowadays treated under different genera once belonged to *Agropyron*, it is impossible to estimate how many accessions of *Agropyron* sp. reported in the present study would need to be transferred to other genera. The largest collections are those of the Western Regional Plant Introduction Station Pullman, USA (26.8%), the Genetic Resources Department Damascus, Syria (15.0%), the USDA Grass Collection in Logan, Utah, USA (10.4%), and the Australian Winter Cereals Collection in Tamworth, Australia (9.2%) (Table 2.19). Eleven species of *Agropyron* are found in genebanks (Table 2.20).

Table 2.19 Nine institutions with more than 80 accessions of *Agropyron*

FAO code	Institute	<i>Agropyron</i> accs.
USA022	Western Regional Plant Introduction Station, USDA-ARS, Washington State University, Pullman, Washington, USA	835
SYR055	Genetic Resources Department, General Commission for Scientific Agricultural Research, Damascus, Syria	466
USA169	Forage and Range Research USDA/ARS, Grass Collection, Logan, Utah, USA	325
AUS003	Australian Winter Cereals Collection, Tamworth, Australia	287
CHN029	Grassland Research Institute, Chinese Academy of Agricultural Sciences, Huhehot, Inner Mongolia, China	183
CHN001	Institute of Crop Germplasm Resources, Chinese Academy of Agricultural Sciences, Beijing, China	155
SWE001	Department of Plant Breeding Research, Swedish University of Agricultural Sciences, Alnarp, Sweden	138
HUN003	Institute for Agrobotany, Tápiósztele, Hungary	87
TUR001	Plant Genetic Resources Department, Aegean Agricultural Research Institute, Izmir, Turkey	87

Table 2.20 Total number of accessions per *Agropyron* species, as reported in the combined database used for the present analysis. *Agropyron* hybrid is a collective designation for named and unnamed hybrids between *Agropyron* species. Additional 1,600 accessions are named only to genus (cf. Table 2.1)

Taxon	Accs.	Genebanks
<i>Agropyron cimmericum</i>	1	1
<i>A. cristatum</i>	973	22
<i>A. dasyanthum</i>	2	1
<i>A. desertorum</i>	212	15
<i>A. fragile</i>	179	14
<i>Agropyron</i> hybrid	38	5
<i>A. michnoi</i>	22	2
<i>A. mongolicum</i>	20	4
<i>A. obtusiusculum</i>	0	1
<i>A. pectinatum</i>	61	10
<i>A. pulcherrimum</i>	1	1
<i>A. pumilum</i>	1	1

Databases. The European *Agropyron* Database is maintained by the Institute of Plant Genetic Resources, Sadovo, Plovdiv, Bulgaria (Shamov 2000), but presently it documents accessions from the hosting institute only. It contains passport data of 56 accessions of 23 cultivars and 33 wild ecotypes, belonging to 13 species.

2.4.11 Other *Triticeae* Species

This material includes a number of *Triticeae* genera not included in the above overviews for individual genera (Sections 2.4.4, 2.4.5, 2.4.6, 2.4.7, 2.4.8, 2.4.9,

2.4.10), represented by a total of 6,307 accessions (this figure also includes the 568 unspecified *Triticeae* accessions that may also belong to genera treated above separately, cf. Table 2.1). These genera have largely only wild species, but some of them include also species cultivated for purposes other than human food (cf. Section 2.3). All hybrid genera except \times *Triticosecale* are also listed here. This group includes material of potential interest for research and wide crossing experiments in the major cereals.

The largest collections of this group of *Triticeae* species are hosted by the Western Regional Plant Introduction Station Pullman, USA (26.7%), the USDA Small Grains Collection in Aberdeen (15.3%), the Department of Plant Breeding Research in Alnarp, Sweden (13.8%), and CIMMYT (7.6%) (Table 2.21).

Among the “minor” genera, the largest are *Leymus* (with 1,291 accessions of 27 species), *Thinopyrum* (740, 11), *Dasyphyrum* (609, 2), \times *Aegilotriticum* (568),

Table 2.21 Eighteen institutions with more than 50 accessions of other *Triticeae*

FAO code	Institute	Other <i>Triticeae</i> accs.
USA022	Western Regional Plant Introduction Station, USDA-ARS, Washington State University, Pullman, Washington, USA	1683
USA029	USDA-ARS, National Small Grains Collection, Aberdeen, Idaho, USA	963
SWE001	Department of Plant Breeding Research, Swedish University of Agricultural Sciences, Alnarp, Sweden	868
MEX002	CIMMYT, El Batán, Mexico	479
USA422	Organic Gardening and Farming Research Center, Kutztown, PA, USA	250
CAN004	Plant Gene Resources of Canada, Saskatoon, Saskatchewan, Canada	242
ITA004	Istituto di Genetica Vegetale, Bari, Italy	166
SYR002	ICARDA, Aleppo, Syria	165
DEU146	Leibniz Institute of Plant Genetics and Crop Plant Research, Gatersleben, Germany	158
CZE122	Crop Research Institute, Prague, Czech Republic	147
BRA015	Centro Nacional de Pesquisa de Trigo, EMBRAPA, Passo Fundo, Brazil	93
CHN001	Institute of Crop Germplasm Resources, Chinese Academy of Agricultural Sciences, Beijing, China	92
SWE002	NordGen, Alnarp, Sweden	87
GRC005	Greek Genebank, Thermi, Thessaloniki, Greece	86
NLD037	Centre for Genetic Resources The Netherlands, Wageningen, The Netherlands	75
CHN029	Grassland Research Institute, Chinese Academy of Agricultural Sciences, Huhehot, Inner Mongolia, China	72
ESP004	Centro de Recursos Fitogenéticos, Alcalá de Henares, Spain	66
ARG037	Estación Experimental Agropecuaria Santa Cruz, INTA, Río Gallegos, Santa Cruz, Argentina	57

Psathyrostachys (412, 4), *Pseudoroegneria* (410, 10), × *Elytricum* (407), *Elytrigia* (307, 5), *Eremopyrum* (218, 4), *Taeniatherum* (194, 1), *Amblyopyrum* (138, 1), × *Elytritricale* (101) (Tables 2.1 and 2.22). The genera *Pascopyrum*, *Hordelymus*, × *Agrotriticum*, *Kengyilia*, × *Tritordeum*, *Heteranthelium*, *Hystrix*, × *Haynaldoticum*, *Henrardia*, *Crithopsis*, × *Elyhordeum*, × *Elytritolops*, × *Elymotriticum*, *Festucopsis*, × *Pseudoroelymus* and × *Hordale* have less than 100 to just 1 accession (listed in

Table 2.22 Total number of accessions per species of other *Triticeae* genera, as reported in the combined database used for the present analysis. *Leymus* hybrid is a collective designation for named and unnamed hybrids between *Leymus* species. Accessions of these genera named only to genus are not listed here (cf. Table 2.1). Hybrid genera without species designations are listed only in Table 2.1

Taxon	Accs.	Genebanks
<i>Amblyopyrum muticum</i>	138	13
<i>Crithopsis delileana</i>	9	5
<i>Dasypyrum breviaristatum</i>	2	2
<i>D. villosum</i>	508	17
<i>Elytrigia caespitosa</i>	6	3
subsp. <i>nodosa</i>	1	1
<i>E. campestris</i>	5	2
<i>E. juneiformis</i>	2	1
<i>E. littoralis</i>	6	1
<i>E. obtusiflora</i>	9	1
<i>Eremopyrum bonaepartis</i>	105	5
<i>E. distans</i>	3	2
<i>E. orientale</i>	22	4
<i>E. triticeum</i>	15	4
× <i>Haynaldoticum hungaricum</i>	7	2
× <i>H. sardoum</i>	1	1
<i>Henrardia persica</i>	7	1
<i>Heteranthelium piliferum</i>	11	3
<i>Hordelymus europaeus</i>	58	11
<i>Kengyilia alata</i>	3	1
<i>K. batalinii</i>	6	3
<i>K. grandiglumis</i>	3	1
<i>K. hirsuta</i>	6	2
<i>K. kokonorica</i>	1	1
<i>K. laxiflora</i>	2	1
<i>K. melanthera</i>	5	1
<i>K. mutica</i>	1	1
<i>K. pulcherrima</i>	1	1
<i>K. rigidula</i>	3	1
<i>K. stenachyra</i>	1	1
<i>K. thoroldiana</i>	1	1

Table 2.22 (continued)

Taxon	Accs.	Genebanks
<i>Leymus akmolinensis</i>	5	3
<i>L. alaicus</i>	7	2
<i>L. ambiguus</i>	5	1
<i>L. angustus</i>	264	8
<i>L. arenarius</i>	128	10
<i>L. chinensis</i>	73	3
<i>L. cinereus</i>	183	3
<i>L. condensatus</i>	4	3
<i>L. coreanus</i>	3	1
<i>L. erianthus</i>	14	4
<i>L. flavescens</i>	1	1
<i>Leymus hybrid</i>	7	1
<i>L. innovatus</i>	14	2
<i>L. karataviensis</i>	1	1
<i>L. karelinii</i>	9	3
<i>L. mollis</i>	111	5
<i>L. multicaulis</i>	20	2
<i>L. ovatus</i>	3	1
<i>L. paboanus</i>	22	8
<i>L. pseudoracemosus</i>	2	1
<i>L. racemosus</i>	69	13
<i>L. ramosus</i>	9	1
<i>L. salina</i>	11	3
<i>L. secalinus</i>	73	5
<i>L. tianschanicus</i>	2	1
<i>L. triticoides</i>	19	2
<i>L. ugamicus</i>	1	1
<i>L. villosissimus</i>	1	1
<i>Pascopyrum smithii</i>	94	13
<i>Psathyrostachys fragilis</i>	16	3
<i>P. huashanica</i>	1	1
<i>P. juncea</i>	295	19
<i>P. lanuginosa</i>	3	1
<i>Pseudoroegneria cognata</i>	6	3
<i>P. divaricata</i>	1	1
<i>P. geniculata</i>	26	3
<i>P. gracillima</i>	2	1
<i>P. kosaninii</i>	1	1
<i>P. libanotica</i>	21	4
<i>P. spicata</i>	173	6
<i>P. stipifolia</i>	15	4
<i>P. strigosa</i>	21	3
<i>P. tauri</i>	30	2
<i>Taeniatherum caput-medusae</i>	47	8
subsp. <i>asper</i>	16	4
subsp. <i>caput-medusae</i>	6	2
subsp. <i>crinitum</i>	62	9

Table 2.22 (continued)

Taxon	Accs.	Genebanks
<i>Thinopyrum curvifolium</i>	2	2
<i>T. distichum</i>	1	1
<i>T. elongatum</i>	176	22
<i>T. gentryi</i>	2	2
<i>T. intermedium</i>	462	22
<i>T. junceum</i>	41	10
<i>T. podperae</i>	1	1
<i>T. ponticum</i>	15	3
<i>T. pungens</i>	7	6
<i>T. pycnanthum</i>	8	5
<i>T. scirpeum</i>	1	1

descending order, for figures cf. Table 2.1), with one or few species each, an exception being *Kengyilia* with 12 spp. in genebanks (Table 2.22).

There are also intergeneric hybrids without a hybrid name on the genus level, i.e., *Elymus repens* \times *Pseudoroegneria spicata* (9 accessions in 3 genebanks), *Elymus repens* \times *Thinopyrum pungens* subsp. *campestre* (2 accessions in 1 genebank), and one trigeneric hybrid, *Triticum* \times *Secale* \times *Thinopyrum*.

2.4.12 *Brachypodium*

This genus is included in this survey because of the importance of *B. distachyon* as a new model species in genomics research. It is, however, not reflected in the above summaries (Tables 2.1, 2.2, and 2.3), since following the circumscription of *Triticeae* adopted here (cf. Chapter 1), it is outside this tribe. Twenty-four collections were identified containing 412 accessions belonging to 10 species (Tables 2.23 and 2.24). Relatively large research collections of *B. distachyon* exist in the Western Regional Plant Introduction Station, USA, and in Aberystwyth, UK (Table 2.23).

Garvin et al. (2008) note that collecting of the available diversity in *Brachypodium* is of high importance, and describe the known distribution. Via the search portal of the Global Biodiversity Information Facility (GBIF 2008), almost 52,000 occurrence records (herbarium and living specimens, floristic observations) for *Brachypodium* were found, among them 1,906 of *B. distachyon* (including its synonyms *Bromus distachyos* and *Trachynia distachya*) of which 1,292 are georeferenced and can be displayed on an interactive map.

Table 2.23 Nine institutions with more than 10 accessions of *Brachypodium*

FAO code	Institute	<i>Brachypodium</i> accs.
USA022	Western Regional Plant Introduction Station, USDA-ARS, Washington State University, Pullman, Washington, USA	190 ¹
GBR	Institute of Biological Sciences, Aberystwyth University, Wales, UK	53 ²
GBR004	Millennium Seed Bank Project, Royal Botanic Gardens, Kew, Wakehurst Place, UK	27
DEU146	Leibniz Institute of Plant Genetics and Crop Plant Research, Gatersleben, Germany	22
CZE082	Oseva Pro Ltd. Grassland Research Station Zubri, Czech Republic	20
HUN003	Institute for Agrobotany, Tápiószéle, Hungary	15
ITA004	Istituto del Germoplasma, Bari, Italy	15
BGR001	Institute for Plant Genetic Resources, Sadovo, Bulgaria	11
TUR001	Plant Genetic Resources Department, Aegean Agricultural Research Institute, Izmir, Turkey	11

¹ Including 27 accessions of *B. distachyon* published by Garvin (2007); the total figure also includes unavailable and historical accessions

² Source: germplasm list published under Brachyomics (2008)

Table 2.24 Total number of accessions per *Brachypodium* taxon, as reported by various sources used in the present analysis

Taxon	Accs.	Genebanks
<i>B. bolusii</i>	2	2
<i>B. distachyon</i>	107	6
<i>B. flexum</i>	2	1
<i>B. japonicum</i>	4	1
<i>B. mexicanum</i>	2	1
<i>B. phoenicoides</i>	31	2
<i>B. pinnatum</i>	125	12
<i>B. retusum</i>	5	1
<i>B. rupestre</i>	16	2
<i>Brachypodium</i> sp.	42	9
<i>B. sylvaticum</i>	76	10
Total	412	24

2.5 Outlook and Conclusions

Compiling a survey on cultivated and useful species of a large and diverse group of plants such as the *Triticeae* is possible only if good and reliable information sources are available. The same is true for the overview of genebank collections holding material of the *Triticeae*. Such endeavours heavily rely on the great efforts put into such compilations by others.

For the cultivated and useful species, two information sources of high quality are available, namely, the Mansfeld Database and the USDA online information. For the overview of plant genetic resources collections, two global information sources exist, based on the same information compiled in the mid 1990s and later updated separately and differently (cf. Section 2.2.1). Research into the actual composition of single *Triticeae* collections, as well as information from more recent overviews for particular genera (e.g., the Crop Strategies) and from European Central Crop Databases shows, however, that the reliability of the figures provided by the two global information systems, the FAO WIEWS database and the Bioversity Directory, is deteriorating (for considerable improvements with respect to the WIEWS database, see, however, footnote 1 in Section 2.2.1). Genebanks discontinue to exist, they are being merged or split, collections are lost or newly collected and acquired, or move from one place to another. All these processes need to be monitored and reflected in these global databases. The global plant genetic resources community, including their customers, i.e., breeders and researchers, need these global information sources to identify collections from where they can acquire plant material for their work.

The compilation of the information on cultivated and useful *Triticeae* species and on the global holdings of *Triticeae* germplasm presented in this chapter can be regarded as a contribution to an integrated information system on *Triticeae* taxonomy and nomenclature, with links to existing living collections of this material.

Acknowledgments The present study could not have been compiled without the invaluable support, contributions and discussions from numerous colleagues from the world of *Triticeae* taxonomy and nomenclature, and from genebank curators and information managers in genetic resources. I am grateful to all of them. Mary Barkworth and Roland von Bothmer provided useful comments on *Triticeae* classification and nomenclature and checked hundreds of names for the synonymy, but also suggested improvements to earlier versions of this manuscript. Laura Morrison commented especially on names in *Triticum* and *Aegilops*. Tom Payne provided information about the Wheat Strategy of the Global Crop Diversity Trust, and Jan Valkoun on the Barley Strategy. John Wiersema supported the study with providing taxonomic and other data from GRIN, and with relevant discussions. Adugna Abdi Woldeamayyat updated me on the holdings of the Institute of Biodiversity Conservation and Research, Addis Ababa, Ethiopia, Iva Faberová provided summary information from the Czech genebank and the European Wheat Database, Edson Jair Iorczeski informed about the *Triticeae* holdings of Brazilian genebanks, Gert Kleijer sent information from the European Triticale Database, Jan Konopka provided data from the Global Barley Register (in progress), Giambattista Polignano updated information about the Bari Genebank in Italy, Marcin Zaczynski sent data from the European *Secale* Database. I am grateful to the anonymous reviewer for corrections and useful suggestions of an earlier version of the manuscript. Catherine Feuillet provided continuous support and encouragement as well as valuable advice for improving and completing the coverage of the chapter. Finally, I wish to thank my colleagues Andreas Börner, Annette Weidner and Benjamin Kilian for helpful discussions and information, and Markus Oppermann for assistance in providing IPK's passport data.

Appendix: Online Databases

Database	URL
Barley Pedigree Catalogue	http://genbank.vurv.cz/barley/pedigree/
Bioversity Directory of Germplasm Collections	http://www.bioversityinternational.org/Information_Sources/Germplasm_Databases/Germplasm_Collection_Directory/index.asp
ECPGR Cereals Network website	http://www.ecpgr.cgiar.org/Networks/Cereals/cereals.htm
EURISCO – European Plant Genetic Resources Search Catalogue	http://eurisco.ecpgr.org/index.php
European Barley Database of the ECPGR	http://barley.ipk-gatersleben.de/ebdb/
European <i>Secale</i> Database of the ECPGR	http://www.ihar.edu.pl/gene_bank/secale/secale.html
European Triticale Database of the ECPGR	http://www.bdn.ch/etdb/
European Wheat Database of the ECPGR	http://genbank.vurv.cz/ewdb/
FAO WIEWS – World Information and Early Warning System	http://apps3.fao.org/wiews/institute_query.htm
GBIS – Genebank Information System, IPK Gatersleben, Germany	http://www.ipk-gatersleben.de/Internet/Infrastruktur/Genbankinfosystem
GBIF – Global Biodiversity Information Facility Data Portal	http://www.gbif.net
GRIN – Genetic Resources Information Network of the USDA	http://www.ars-grin.gov/npgs/acc/acc_queries.html
GRIN Canada – Genetic Resources Information Network of Plant Gene Resources of Canada	http://pgrc3.agr.gc.ca/search_grinca-recherche_rircg_e.html
Mansfeld's World Database of Agricultural and Horticultural Crops	http://mansfeld.ipk-gatersleben.de/mansfeld
SINGER – System-wide Information Network for Genetic Resources	http://singer.cgiar.org/
Taxonomic Nomenclature Checker – GRIN	http://pgrdoc.bioversity.cgiar.org/taxcheck/grin/index.html
USDA-ARS – World Economic Plants Report	http://www.ars-grin.gov/cgi-bin/npgs/html/taxeconlist.pl
Wheat Genetic and Genomic Resources Center (WGGRC), Kansas State University	http://www.k-state.edu/wgrc/
Wheat Pedigree Catalogue	http://genbank.vurv.cz/wheat/pedigree/

References

- Arias, G., A. Reiner, A. Penger and A. Mangstl (1983) *Directory of barley cultivars and lines*. Ulmer, Stuttgart.
- Barkworth, M.E., K.M. Capels, S. Long, L.K. Anderton and M.B. Piep (Eds.) (2007) *Flora of North America North of Mexico*. Vol. 24. Magnoliophyta: Commeliniferae (in part): Poaceae, part 1. Oxford University Press, New York, Oxford, xxviii + 911 pp.

- Baum, B.R. (1995) Triticale Register, an on-line register. *Acta Hort.* 413, 163–168.
- Baum, B.R., L.G. Bailey, J. Nugent, W.G. Royds and R. Elvidge (1990) The Triticale Cultivars International Register Computer System, a flexible information system suitable for international registration. *Taxon* 39, 9–15.
- Baum, B.R., L.G. Bailey and B.K. Thompson (1985) *Barley register*. Canada Gov. Publ. Centre. Cat. No. A53-1783/1985. Biosystematics Res. Institute, Ottawa, Ontario, Canada.
- Baumer, M. and R. Cais (2000) *Abstammung der Gerstensorten*. Bayer. Landesanst. Bodenkultur u. Pflanzenbau, Freising, 109 pp. Also online under: http://www.lfl.bayern.de/ipz/gerste/09740/linkurl_0_9.pdf
- Bioversity (2007) Bioversity Directory of Germplasm Collections, accessed November/December 2007.
- Bockelman, H. (1996) Developing a barley core subset in the USDA-ARS National Small Grains Collection. In: A. Slinkard, G. Scoles and B. Rosnagel (Eds.), *Proc. Vth Internat. Oats Conf. & VIIth Internat. Barley Genetics Sympos.*, Saskatoon, Canada, p. 154.
- Bor, N.L. (1968) *Gramineae*. In: C.C. Townsend, E. Guest and A. Al-Rawi (Eds.), *Flora of Iraq*. Vol. 9. Ministry of Agriculture, Baghdad, vi + 588 pp.
- Börner, A. and A.J. Worland (1995) List of cytogenetic tester stocks – the reply of the questionnaire sent out with the 1991 EWAC-Newsletter. *European Wheat Aneuploid Co-operative Newsletter* 1995, 180–183.
- Bothmer, R. von, N. Jacobsen, C. Baden, R.B. Jørgensen and I. Linde-Laursen (1995) An ecogeographical study of the genus *Hordeum*. *Systematic and Ecogeographical Studies on Crop Gene pools* 7. IPGRI, Rome.
- Bothmer, R. von, O. Seberg and N. Jacobsen (1992) Genetic resources in the *Triticeae*. *Hereditas* 116, 141–150.
- Brachyomics (2008) Germplasm [University of Wales, Aberystwyth]. <http://www.aber.ac.uk/plantpathol/germplasm.htm>. Accessed Jul 2008.
- Cadle, M.M., T.D. Murray and S.S. Jones (1997) Identification of resistance to *Pseudocercospora herpotrichoides* in *Triticum monococcum*. *Plant Dis.* 81, 1181–1186.
- Ceccarelli, S., C.A. Cenci and E. Piano (1975) Selezione per più caratteri in un programma di miglioramento genetico del *Brachypodium pinnatum* P. B. *Genet. Agrar.* 29, 95–110.
- Chabane, J.C.K. and J. Valkoun (2004) Characterisation of genetic diversity in ICARDA core collection of cultivated barley (*Hordeum vulgare* L.). *Czech J. Genet. Plant Breed.* 40, 134–136.
- Chen, Q., R.H. Zhou, L.H. Li, X.Q. Li, X.M. Yang and Y.S. Dong (1988) First intergeneric hybrid between *Triticum aestivum* and *Psathyrostachys juncea*. *Kexue Tongbao* (Beijing) 33, 2071–2074. (cited after Wei et al. 1997)
- CIMMYT (2007) Global Strategy for the *ex situ* conservation with enhanced access to wheat, rye and triticale genetic resources. International Maize and Wheat Improvement Center (CIMMYT), Mexico. <http://www.croptrust.org/documents/web/Wheat-Strategy-FINAL-20Sep07.pdf>. Accessed Jan 2008.
- Cubadda, R. and E. Marconi (1996) Technological and nutritional aspects in emmer and spelt. In: S. Padulosi, K. Hammer and J. Heller (Eds.), *Hulled wheats. Promoting the conservation and use of underutilized and neglected crops*. 4. International Plant Genetic Resources Institute, Rome, Italy, pp. 203–211.
- Darvey, N.L., G. Oettler and W.H. Pfeiffer (1996) Establishment of a network for genetic resources. In: H. Guedes-Pinto, N. Darvey and V.P. Carnide (Eds.), *Triticale: Today and Tomorrow. Developments in Plant Breeding* 5. Kluwer, pp. 269–274.
- Dorofeev, V.F., A.A. Filatenko, E.F. Migushova, R.A. Udaczin and M.M. Jakubziner (1979) *Wheat*. In: V.F. Dorofeev and O.N. Korovina (Eds.), *Flora of Cultivated Plants*. vol. 1. Leningrad, Russia, Kolos, 346 pp. (in Russian)
- Dosba, F. and G. Dossinault (1973) Resistance to eyespot (*Cercospora herpotrichoides*) introduced to bread wheat from *Aegilops ventricosa*. In: *Proc. Int. Wheat Genet. Symp.* 4, pp. 409–413.

- Draper J., L.A.J. Mur, G. Jenkins, G.C. Ghosh-Biswas, P. Bablak, R. Hasterok and A.P.M. Routledge (2001) *Brachypodium distachyon*. A new model system for functional genomics in grasses. *Plant Physiol.* 127, 1539–1555.
- EBDB (2008) European Barley Database of the European Cooperative Programme for Plant Genetic Resources. IPK Gatersleben, Germany. Accessed Feb 2008.
- Enneking, D. and H. Knüpfper (2001) The European Barley Database. In: H. Knüpfper, R. von Bothmer, M. Ambrose, R. Ellis, A.M. Stanca, D. Enneking, L. Maggioni and E. Lipman (Compilers), *Report of a Working Group on Barley, Sixth Meeting, 3 December 2000, Salsomaggiore, Italy*. International Plant Genetic Resources Institute, Rome, Italy, pp. 50–62.
- ETDB (2008) European Triticale Database. Agroscope Changins-Wädenswil Research Station (ACW), Nyon, Switzerland. Accessed Mar 2008.
- EURISCO (2008) European Plant Genetic Resources Search Catalogue. Accessed Feb 2008.
- European Secale Database (2008) European *Secale* Database. IHAR, Radzików, Poland. Accessed Mar 2008.
- EWDB (2008) European Wheat Database. Crop Research Institute, Prague, Czech Republic. Accessed Mar 2008.
- FAO (1996) *Report on the State of the World's Plant Genetic Resources for Food and Agriculture, prepared for the International Technical Conference on Plant Genetic Resources Leipzig, Germany, 17–23 June 1996*. Food and Agriculture Organization of the United Nations, Rome, 81 pp. www.fao.org/AG/AGp/agps/Pgrfa/pdf/swrshr_e.pdf, accessed January 2008.
- FAO (1997) *The State of the World's Plant Genetic Resources for Food and Agriculture*. Food and Agriculture Organization of the United Nations, Rome, 511 pp. Available from: <http://www.fao.org/ag/agp/AGPS/Pgrfa/pdf/swrfull.pdf>. Accessed Jan 2008.
- FAO (2007) WIEWS – World Information and Early Warning System. Accessed Nov/Dec 2007.
- Garvin, D. (2007) The USDA-ARS NPGS *Brachypodium distachyon* collection. Available from: http://www.brachypodium.org/files/File/USDA-ARS_Brachypodium_Collection.pdf. Accessed 16 Jul 2008.
- Garvin, D.F., Y.-Q. Gu, R. Hasterok, S.P. Hazen, G. Jenkins, T.C. Mockler, L.A.J. Mur and J.P. Vogel (2008) Development of genetic and genomic research resources for *Brachypodium distachyon*, a new model system for grass crop research. *Crop Sci.* 48(S1), S69–S84.
- Gass, T., W. Podyma, J. Puchalski and S.A. Eberhart, compilers (1998) *Challenges in Rye Germplasm Conservation*. International Plant Genetic Resources Institute, Rome, Italy.
- GBIF (2008) Global Biodiversity Information Facility Data Portal. Accessed Jul 2008.
- GBIS (2008) Genebank Information System of the Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Gatersleben, Germany. Accessed Feb 2008.
- GRIN (2008) Genetic Resources Information Network of the USDA. Accessed Feb 2008.
- GRIN Canada (2008) Genetic Resources Information Network of Plant Gene Resources of Canada. Accessed Feb 2008.
- Hammer, K. (1980) Vorarbeiten zur monographischen Darstellung von Wildpflanzensortimenten: *Aegilops* L. *Kulturpflanze* 28, 33–180.
- Hammer, K., E. Skolimowska and H. Knüpfper (1987) Vorarbeiten zur monographischen Darstellung von Wildpflanzensortimenten: *Secale* L. *Kulturpflanze* 35, 135–177.
- Hanelt, P. (2001) *Gramineae (Poaceae)*. Trib. *Triticeae*. In: Hanelt, P. and Institute of Plant Genetics and Crop Plant Research (Eds.), *Mansfeld's Encyclopedia of Agricultural and Horticultural Crops*. vol. 5. Springer, Berlin, Heidelberg, New York etc., pp. 2542–2594.
- Hao, C.Y., X.Y. Zhang, L.F. Wang, Y.S. Dong, X.W. Shang and J.Z. Jia (2006) Genetic diversity and core collection evaluations in common wheat germplasm from the North-western Spring Wheat Region in China. *Mol. Breeding* 17, 69–77.
- Hasterok, R., J. Draper, and G. Jenkins (2004) Laying the cytotoxic foundations of a new model grass, *Brachypodium distachyon* (L.) Beauv. *Chromosome Res.* 12, 397–403.

- Hintum, Th.J.L. van and F. Menting (2000) Barley genetic resources conservation – now and forever. In: S. Logue (Ed.), *Barley Genetics VIII*. vol. I, Adelaide Univ., Glen Osmond, South Australia, pp. 13–20.
- Hintum, Th.J.L. van and F. Menting (2003) Diversity in *ex situ* genebank collections of barley. In: R. von Bothmer, Th.J.L. van Hintum, H. Knüpfper and K. Sato (Eds.), *Diversity in Barley* (*Hordeum vulgare*). Elsevier Science B.V., Amsterdam, The Netherlands, pp. 247–257.
- Hodgkin, T., Y.J. Adham and K.S. Powell (1992) A preliminary survey of wild *Triticum* and *Aegilops* species in the world's genebanks. *Hereditas* 116, 155–162.
- Igartua, E., M.P. Gracia, J.M. Lasa, B. Medina, J.L. Molina-Cano, J.L. Montoya and I. Romagosa (1998) The Spanish barley core collection. *Genetic Resour. Crop Evol.* 45, 475–481.
- International Brachypodium Initiative (2008) *Brachypodium* Sequence Resource Gateway. The *Brachypodium distachyon* Information Resource. Available from: <http://www.brachypodium.org/>. Accessed 16 Jul 2008.
- Kimber, G. and M. Feldman (1987) *Wild Wheat. An Introduction*. Coll. Agric. Univ. Missouri-Columbia, Spec. Rep. 353.
- Kleijer, G. (2007) The Triticale Central Crop Database. In: G. Kleijer, R. Häner and H. Knüpfper (Compilers), *Triticale and Rye Genetic Resources in Europe*. Bioversity International, Rome, Italy, pp. 4–5.
- Kleijer, G., R. Häner and H. Knüpfper, compilers (2007) *Triticale and Rye Genetic Resources in Europe*. Bioversity International, Rome, Italy, iv + 15 pp.
- Knüpfper, H. (1988) The European Barley Database of the ECP/GR: An introduction. *Kulturpflanze* 36, 135–162.
- Knüpfper, H. (2008) Barley Working Group, Summary of country reports. In: Maggioni et al. (2008), unpublished.
- Knüpfper, H. and Th.J.L. van Hintum (1995) The Barley Core Collection – an international effort. In: T. Hodgkin, A.H.D. Brown, Th.J.L. van Hintum and E.A.V. Morales (Eds.), *Core Collections of Plant Genetic Resources*. John Wiley & Sons, Chichester etc., pp. 171–178.
- Knüpfper, H. and Th. van Hintum (2003) Summarised diversity – the Barley Core Collection. In: R. von Bothmer, Th. van Hintum, H. Knüpfper and K. Sato (Eds.), *Diversity in Barley* (*Hordeum vulgare*). Elsevier Science B.V., Amsterdam, The Netherlands, pp. 259–267.
- Kobiljski, B., S. Quarrie, S. Denčić, J. Kirby and M. Ivegeš (2002) Genetic diversity of the Novi Sad wheat core collection revealed by microsatellites. *Cell. Mol. Biol. Lett.* 7, 685–694.
- Konopka, J. and J. Valkoun (2005) Global database of wheat wild relatives. *Czech J. Genet. Plant Breed.* 41, 251.
- Kovács, G., A.F. Bálint and J. Sutka (2002) Activity of the Cereal Genebank and the perspectives of diploid wheat research in Martonvásár. In: W. Świ cicki, B. Naganowska and B. Wolko (Eds.), *Broad variation and precise characterization limitation for the future*, pp. 45–48.
- Lipman, E., L. Maggioni, H. Knüpfper, R. Ellis, J.M. Leggett, G. Kleijer, I. Faberová and A. Le Blanc, compilers (2005) *Cereal Genetic Resources in Europe*. International Plant Genetic Resources Institute, Rome, Italy, 318 pp.
- Löve, Á. (1984) Conspectus of the *Triticeae*. *Feddes Repert.* 95, 425–521.
- Lundqvist, U. (2007) Overall coordinators report. *Barley Genet. Newsl.* 37, 153–187.
- Lundqvist, U. and M. Huldén (2005) International Barley Genes and Barley Genetic Stocks database. In: E. Lipman; L. Maggioni, H. Knüpfper, R. Ellis, J.M. Leggett, G. Kleijer, I. Faberová and A. Le Blanc (Eds.), *Cereal Genetic Resources in Europe*. International Plant Genetic Resources Institute, Rome, Italy, pp. 128–129.
- Maggioni, L., A. Katsiotis, H. Knüpfper and G. Kleijer, compilers (2008) *Report of a Cereals Network meeting*, Foça, Izmir, Turkey, 21–24 April 2008. Bioversity International, Rome.

- Unpublished draft. Final version to be published under <http://www.ecpgr.cgiar.org/Networks/Cereals/cereals.htm>
- Malosetti, M., T. Abadie and S. German (2000) Comparing strategies for selecting a core subset for the Uruguayan barley collection. *FAO/IPGRI Plant Genet. Resour. Newsl.* 121, 20–26.
- Mansfeld Database (2008) Mansfeld's World Database of Agricultural and Horticultural Crops. Accessed Dec 2007, Jan 2008.
- Martín, A., C. Martínez-Araque, D. Rubiales and J. Ballesteros (1996) Tritordeum: triticales's new brother cereal. In: H. Guedes-Pinto, N. Darvey and V.P. Carnide (Eds.), *Triticale: Today and Tomorrow. Developments in Plant Breeding* 5. Kluwer, pp. 57–72.
- Martynov, S.P., T.V. Dobrotvorskaya, L. Dotlačil, Z. Stehno, I. Faberová and I. Bareš (2003) Genealogical approach to the formation of the winter wheat core collection. *Genetika* 39, 1091–1098. (also in: *Russ. J. Genetics* 39, 917–923).
- Martynov, S.P., T.V. Dobrotvorskaya, I. Hon and I. Faberová (2006a) Wheat Pedigree and Identified Alleles of Genes On Line. Last updated 8 September 2006.
- Martynov, S.P., T.V. Dobrotvorskaya, I. Hon and I. Faberová (2006b) Barley Pedigree Catalogue on Line.
- Martynov, S.P., T.V. Dobrotvorskaya, Z. Stehno, L. Dotlačil, I. Faberová and V. Holubec (1992) *Genealogies and Gene Alleles Identified in 31000 Cultivars and Lines of Wheat*. Res. Inst. Crop Product. Press, Prague.
- Maschka, R. (1995) Die Identifizierung von Duplikaten in Genbankkollektionen im Rahmen des Informationssystems Genetischer Ressourcen in Deutschland. Diploma Thesis, Univ. Kassel Gesamthochschule.
- Maxted, N., K. White, J. Valkoun, J. Konopka and S. Hargreaves (2008) Towards a conservation strategy for *Aegilops* species. *Plant Genet. Resour.* 6, 126–141.
- McNeill, J., F.R. Barrie, H.M. Burdet, V. Demoulin, D.L. Hawksworth, K. Marhold, D.H. Nicolson, J. Prado, P.C. Silva, J.E. Skog, J.H. Wiersema and N.J. Turland (2006) *International Code of Botanical Nomenclature (Vienna Code)*. Regnum Vegetabile 146. A.R.G. Gantner Verlag KG.
- Müntzing, A. (1974) Historical review of the development of triticales. In: R. MacIntyre and M. Campbell (Eds.), *Triticale: Proc. Internat. Symposium*. International Development Research Centre, Ottawa, pp. 13–30.
- Pace, C. de, R. Paolini, G.T. Scarascia Mugnozza, C.O. Qualset and V. Delre (1990) Evaluation and utilization of *Dasypyrum villosum* as a genetic resource for wheat improvement. In: J.P. Srivastava and A.B. Damania (Eds.), *Wheat genetic resources: meeting diverse needs*. John Wiley and Sons, Chichester (UK), pp. 279–288.
- Perrino, P., G. Laghetti, L.F. D'Antuono, M. Al Ajlouni, M. Kanbertay, A.T. Szabó and K. Hammer (1996) Ecogeographical distribution of hulled wheat species. In: S. Padulosi, K. Hammer and J. Heller (Eds.), *Hulled wheats. Promoting the conservation and use of under-utilized and neglected crops*. 4. International Plant Genetic Resources Institute, Rome, Italy, pp. 101–119.
- Plourde, A., G. Fedak, C.A. St-Pierre and A. Comeau (1990) A novel intergeneric hybrid in the *Triticeae*: *Triticum aestivum* × *Psathyrostachys juncea*. *Theor. Appl. Genet.* 79, 45–48. (cited after Wei et al. 1997)
- Podyma, W. (1998) Rye genetic resources in European genebanks. In: T. Gass, W. Podyma, J. Puchalski and S.A. Eberhart (Compilers), *Challenges in Rye Germplasm Conservation*. International Plant Genetic Resources Institute, Rome, Italy, pp. 87–92.
- Rabinovich, S.V. (1998) Importance of wheat-rye translocations for breeding modern cultivars of *Triticum aestivum* L. *Euphytica* 100, 323–340.
- Schubert, V. (2000) Die Sammlung und Charakterisierung pflanzlicher genetischer Ressourcen und ihre Nutzung im Rahmen der Züchtungsforschung bei Weizen. *Diss. Botanicae* 352. J. Cramer Verlagbuchhandlung, Berlin, Stuttgart, 116 pp.

- Shamov, D.P. (2000) European *Agropyron* Database. In: L. Maggioni, P. Marum, N.R. Sackville Hamilton, M. Huldén and E. Lipman (Compilers), *Report of a Working Group on Forages, Seventh Meeting, 18–20 November 1999, Elvas, Portugal*. International Plant Genetic Resources Institute, Rome, Italy, p. 59.
- SINGER (2008) System-wide Information Network for Genetic Resources. Accessed Feb 2008.
- Spagnoletti Zeuli, P.L. and C.O. Qualset (1993) Evaluation of five strategies for obtaining a core subset from a large genetic resource collection of durum wheat. *Theor. Appl. Genet.* 87, 295–304.
- Spagnoletti Zeuli, P.L.S. and C.O. Qualset (1995). The durum wheat core collection and the plant breeder. In: T. Hodgkin, A.H.D. Brown, Th.J.L. van Hintum and E.A.V. Morales (Eds.), *Core Collections of Plant Genetic Resources*. John Wiley and Sons, UK, pp. 213–228.
- Stehno, Z., I. Faberová, L. Dotlačil, S. Martynov and T. Dobrotvorskaya (2006) Genealogical analysis in the Czech spring wheat collection and its use for the creation of core collection. *Czech J. Genet. Plant Breed.* 42, 117–125.
- Szigat, G., M. Herrmann and H. Rapke (1998) Integration von Bastarden mit der Wildgerste *Hordeum bulbosum* in den Zuchtprozeß von Wintergerste. *Schriften Genet. Ressourcen* 8, 267–271.
- Tanaka, T. (1976) *Tanaka's Cyclopaedia of Edible Plants of the World*. Keigohu Publ. Co., Tokyo.
- TNC-GRIN (2008) Taxonomic Nomenclature Checker - GRIN [online]. Accessed Jan/Feb 2008.
- Tzvelev, N.N. (1976) *Zlaki SSSR. Poaceae URSS*. Nauka Leningrad. 788 pp. (in Russian)
- USDA-ARS (2008) World Economic Plants Report from USDA, ARS, National Genetic Resources Program. Germplasm Resources Information Network (GRIN). Accessed 20 Feb 2008.
- Valkoun, J. (compiler) (2008) *Global Strategy for the ex situ conservation and use of barley germplasm*. Unpublished draft.
- Wei, J.-Z., W.F. Campbell and R.R.-C. Wang (1997) Genetic variability in Russian wildrye (*Psathyrostachys juncea*) assessed by RAPD. *Genet. Resour. Crop Evol.* 44, 117–125.
- Yahiaoui, S., E. Igartua, M. Moralejo, L. Ramsay, J.L. Molina-Cano, F.J. Ciudad, J.M. Lasa, M.P. Gracia and A.M. Casas (2008) Patterns of genetic and eco-geographical diversity in Spanish barleys. *Theor. Appl. Genet.* 11, 271–282.
- Zaczyński, M. (2007) The Rye Central Crop Database. In: G. Kleijer, R. Häner and H. Knüpffer (Compilers), *Triticale and Rye Genetic Resources in Europe*. Bioversity International, Rome, Italy, pp. 2–4.
- Zaharieva, M., S. Santoni and J. David (2001) Use of RFLP markers to study genetic diversity and to build a core-collection of the wild wheat relative *Ae. geniculata* Roth (= *Ae. ovata* L.). *Genet. Sel. Evol.* 33(Suppl. 1), S269–S288.
- Zeven, A.C. and N.C. Zeven-Hissink (1976) *Genealogies of 14000 Wheat Cultivars*. Netherlands Cereal Centre, Wageningen; International Maize and Wheat Improvement Centre, Mexico.

Chapter 3

Domestication of the *Triticeae* in the Fertile Crescent

Benjamin Kilian, Hakan Özkan, Carlo Pozzi, and Francesco Salamini

Abstract About 12,000 years ago, humans began the transition from hunter-gathering to a sedentary, agriculture-based society. From its origins in the Fertile Crescent, farming expanded throughout Europe, Asia and Africa, together with various domesticated plants and animals. Where, how and why agriculture originated is still debated. Progress has been made in understanding plant domestication in the last few years. New insights were obtained mainly due to (I) the use of comprehensive germplasm collections covering the whole distribution area for each species; (II) the comparison of many wild and domesticated accessions for each species; (III) the identification of the wild progenitor in the wild gene pool and its comparison with domesticate descendants; (IV) the use of molecular fingerprinting techniques at many loci and the access to new generation high-throughput sequencing technologies; (V) the identification and cloning of genes involved in domestication; and (VI) excavation campaigns.

This chapter reviews the recent knowledge on wheat, barley and rye domestication in the Fertile Crescent and covers several issues concerning the molecular knowledge of the effects induced by domestication and breeding of these crops.

Cereals provide more than 50% of the worldwide crop production and are important renewable resources for food, feed, and industrial materials (faostat.fao.org). The *Triticeae* tribe within the *Pooideae* subfamily of the grass family *Poaceae* includes the crop genera *Triticum* (wheat), *Hordeum* (barley) and *Secale* (rye). Wheat is the primary cereal of temperate regions and the staple food for about 40% of the world's population. Globally, wheat is the second most widely grown crop, just recently

B. Kilian (✉)

Institute of Botany III, Heinrich-Heine-Universität Düsseldorf, Universitätsstrasse 1, 40225 Düsseldorf, Germany; Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Genebank/Genome Diversity, Corrensstrasse 3, 06466 Gatersleben, Germany; Max Planck Institute for Plant Breeding Research, Department of Plant Breeding and Genetics, Carl-von-Linné-Weg 10, 50829 Köln, Germany
e-mail: kilian@ipk-gatersleben.de

superseded by maize, while barley ranks fourth after maize, wheat and rice (faostat.fao.org; www.croptrust.org). Wheat and barley are the most important staple crops of Europe and of the western part of Asia. Wheat is mainly used for bread and pasta, barley as fodder and for brewing beer, and rye is used for bread and fodder. Human history is closely interwoven with these three staple crops, because wheats and barley (and possibly also rye) belong to the Neolithic founder crops that built western agriculture.

3.1 Origins of Cultivated Plants and Agriculture – A Brief Historical Overview

The origin of cultivated plants and their domestication have been of relevant interest beginning with the essays of Alexander von Humboldt (*Essai sur la géographie des plantes*”, Humboldt von 1806; Fiedler and Leitner 2000), Charles Darwin (*The origin of species*”, Darwin 1859 and *The variation of animals and plants under domestication*”, Darwin 1868), and Alphonse de Candolle (*Origine de plantes cultivées*”, Candolle de 1882; Damania 1998).

De Candolle studied biogeography of cultivated crops and indicated three regions where plant domestication may have taken place: Southwest Asia, China and Tropical Asia. According to him, historic events such as glaciations and domestication had paramount importances for crop plant distribution (Candolle de 1882). He pointed out the archaeological, botanical, historical and linguistic evidence that could help to study the origin of plant domestication (Gepts 1998).

In 1926, Nikolay Ivanovich Vavilov published his book *Centers of origin of cultivated plants*” (Vavilov 1926). Vavilov noted that *“the entire varietal and racial diversity of the field and vegetable crops is concentrated in mountainous districts”*. Vavilov summarized all his work on diversity in 1935 in *“The phytogeographical basis for plant breeding”* in which he describes eight centers, including a Mediterranean Center where wheats, barleys, vegetables and fruits originated (Vavilov 1992; Hawkes 1998). Vavilov undertook more than one hundred collecting missions and expeditions, the results of which are presented in the comprehensive collection by the All-Russian Scientific Research Institute of Plant Industry (VIR) (www.vir.nw.ru). Two years later, the archaeologist and philosopher Vere Gordon Childe presented his *“Oasis Theory”* which proposed that agriculture began in the Near East when the climate changed at the end of the last glacial period, a process that he termed *“Neolithic Revolution”* (Childe 1928, 1936; Harris 1998).

Subsequent work by Robert Braidwood who excavated Jarmo (Braidwood and Braidwood 1950) and Cayönü (Braidwood et al. 1969) led to the suggestion that agriculture began in the *“hilly flanks of breasted’s ‘fertile Crescent’”* (Braidwood and Braidwood 1950; Braidwood 1972; Braidwood et al. 1983).

The term “*Fertile Crescent*” stems in turn from James Henry Breasted (Breasted 1938; Braidwood 1972).

Archaeological evidence is, however, not sufficient and contributions from related research fields, like archaeobotany and botany, have been a valuable contribution to our knowledge on the origin of agriculture (Harlan and Zohary 1966; Harlan 1971, 1975, 1995; Hillman and Davies 1990; Nesbitt 1995, 2002; Nesbitt and Samuel 1996; Willcox 1996, 2005; Zohary 1999; Hillman 2000; Tanno and Willcox 2006).

For more than two decades the use of molecular markers has provided new information on genetic diversity of crop plants; in relation to wild relatives, centers of domestication, time frame of the domestication process and specific alleles supporting domesticated traits. The connection between molecular markers and domestication geography was significantly highlighted by Heun et al. (1997) who located the origin of einkorn wheat domestication to the Karacadag mountain area in South East Turkey. Other important contributions using different molecular markers for other species followed: barley (Badr et al. 2000; Kilian et al. 2006; Morrell and Clegg 2007); einkorn (Kilian et al. 2007b); emmer (Ozkan et al. 2002, 2005; Mori et al. 2003; Luo et al. 2007); maize (Wright et al. 2005); rice (Londo et al. 2006) and sorghum (Hamblin et al. 2006).

3.2 Evolution and Domestication of *Triticeae*

Western agriculture originated in the Fertile Crescent after the last ice age, in aceramic Pre-Pottery Neolithic (PPN) from about 12,000 to 9,500 years ago (Zohary and Hopf 2000; Nesbitt 2002; Salamini et al. 2002). Archaeological evidence revealed the occurrence of plant remains at different excavation sites, in different stratigraphic layers that were analyzed and radiocarbon dated (Hillman 2000). On the other hand, phytogeographical, botanical and genetical studies have identified the wild progenitors of crop plants, their distribution and their significant morphological and genetical differences (Zohary and Hopf 2000).

It is accepted today that Fertile Crescent agriculture originated in a “*core area*” in South East Turkey, where the distribution of wild forms are molecularly and cytologically closely related to the founder crops (Table 3.1) (Lev-Yadun et al. 2000; Abbo et al. 2006). From there, farming spread throughout Europe, Asia and Africa (Ammerman and Cavalli-Sforza 1984; Nesbitt 2002). The domestication process lasted up to one millennium in the region (Tanno and Willcox 2006), always preceded by cultivation of wild populations before domestication (Weiss et al. 2006).

Wild relatives differ from their crop descendants (Table 3.1) for several phenotypic characteristics, collectively referred to as the “domestication syndrome” (Hammer 1984; Salamini et al. 2002; Pozzi and Salamini 2007). To classify wild and domesticated materials as separate species is unjustified,

Table 3.1 The founder crops of Neolithic agriculture and their wild progenitors

Name	Wild progenitor	Domesticated form
Einkorn wheat ^a	<i>Triticum boeoticum</i>	<i>T. monococcum</i>
Emmer wheat ^b	<i>Triticum dicoccoides</i>	<i>T. dicoccum</i>
Rye	<i>Secale vavilovii</i>	<i>S. cereale</i>
Barley	<i>Hordeum spontaneum</i>	<i>H. vulgare</i>
Lentil	<i>Lens orientalis</i>	<i>L. culinaris</i>
Pea	<i>Pisum humile</i>	<i>P. sativum</i>
Chickpea	<i>Cicer reticulatum</i>	<i>C. arietinum</i>
Bitter vetch	<i>Vicia ervilia</i>	<i>V. ervilia</i>
Flax	<i>Linum bienne</i>	<i>L. usitatissimum</i>

^a It is still unclear yet if *Triticum urartu* was also collected, due to the fact that seeds of *T. urartu* and *T. boeoticum* cannot be distinguished.

^b Evidence supports that wild *Triticum araraticum* was the progenitor of *T. timopheevii*, but distinguishing plant remains of wild *T. dicoccoides* from those of wild *T. araraticum*, as well the two domesticates, is almost impossible.

because usually their crossing progenies are fertile. However it is common to refer to them, at least when describing archaeological and genetical events, as if they are different species, a formality which simplifies the discussion of domestication-related issues. Over the years, taxonomical classifications were developed by geneticists. For wheat, the latest comprehensive, systematic overview was completed in 1979 by Dorofeev and colleagues. This work was published in Russian and was therefore not recognized worldwide, but the translation into English is currently underway and may become a standard for wheat classification. In this chapter, the nomenclature and the genome formula given for *Triticum* by Dorofeev et al. (1979) and the *Aegilops* nomenclature based on van Slageren (1994) is followed. In general, the Zohary and Hopf (2000) classification is accepted, with few modifications where necessary.

Archaeological evidence indicates that plant remains of nine domesticated species appear almost always together at common sites and times. It is therefore assumed that these species have been domesticated together as a “*founder package*” (Lev-Yadun et al. 2000). The wild and domesticated species of the Neolithic founder package are shown in Table 3.1.

Molecular results, mainly concerning genome-wide measures of genetic similarity have traced the origins of domesticated cereals to wild populations of grasses that are still present in the Fertile Crescent (Heun et al. 1997; Ozkan et al. 2005; Luo et al. 2007; Kilian et al. 2007b).

3.2.1 Wheat Evolution and Domestication

Studies of wheat evolution have attracted large attention over the past 100 years. Sakamura (1918), Sax and Sax (1924) and Kihara (1924) used cytogenetic methods and recognized that wheat species fall into three groups based upon

their ploidy level: (i) diploid $2n = 14$ = einkorn wheat; (ii) tetraploid $4n = 28$ = emmer wheats; (iii) hexaploid $6n = 42$ = bread wheats. Bread wheat has no direct hexaploid wild progenitor and possesses three sets of homoeologous chromosomes, designated as A^uA^uBBDD , whose origins have differing degrees of certainty. The *D* chromosomes stem from wild diploid *Aegilops tauschii* through allopolyploidization with the wild A^uA^uBB tetraploid *T. dicoccoides*. The A^u and *B* chromosomes derive from the hybridization between the wild A^uA^u diploid *T. urartu* and a wild diploid *B* genome donor, frequently reported to belong to the *Sitopsis* section of *Aegilops* (which includes 5 species). Brandolini et al. (2006) quantified the genetic relationships among *A* genomes of wheats by AFLP fingerprinting. The results indicate that the *T. urartu* genome, compared to *T. boeoticum*, is about 20% closer to the *A* genomes of polyploid wheats.

The tetraploid A^uA^uBB and A^uA^uGG wheats originated through independent allopolyploidization events between two wild diploid grasses (summarized in Kilian et al. 2007a). Strong evidence points to the wild outcrossing *Aegilops speltoides* (SS) (or a genotype similar to it) as the female parent of tetraploid wheats and to wild *T. urartu* (A^uA^u) as the male parent (Dvorak and Zhang 1990; Huang et al. 2002; Zhang et al. 2002; Kilian et al. 2007a). The hybridization which generated the A^uA^uBB wheats may have taken place between 0.25 and 1.3 MYA (Mori et al. 1995; Huang et al. 2002; Dvorak and Akhunov 2005), while the event that led to the A^uA^uGG wheats likely occurred later (Huang et al. 2002). The evolutionary relationships between wheats, with different ploidy levels, are shown in the Supplementary figure S2 of Kilian et al. (2007a).

3.2.1.1 Diploid Wheats

Two wild diploid *Triticum* species are recognized: *T. boeoticum* (A^bA^b) and *T. urartu* (A^uA^u). They are separated by crossing barriers (Johnson and Dhaliwal 1976), differ in plant morphology (Gandilian 1972; Dorofeev et al. 1979) and at biochemical and molecular marker loci (Johnson 1975; Dvorak et al. 1998a; Kilian et al. 2007b). The diploid einkorn wheat *T. monococcum* was among the first crops domesticated in the Fertile Crescent starting from the wild progenitor *T. boeoticum*. Domestication was located to the geographic area of the volcanic Karacadag mountain range in South East Turkey (Heun et al. 1997). The earliest archaeological records from domesticated einkorn are described from Abu Hureyra (Hillman et al. 1989), Cayönü (van Zeist and de Roller 1991–2), and Nevali Cori (Pasternak 1998). Einkorn was the staple crop of the Shumer populations and has been found in the excavated layers of Troy (Nesbitt and Samuel 1996); currently einkorn is a relict crop. In fact, during the last 5,000 years einkorn was largely abandoned and replaced by tetraploid and hexaploid wheats. Thus, domesticated einkorn germplasm is devoid of modern breeding bottlenecks. Therefore, Kilian et al. (2007b) reasoned that extensive sampling of genetic diversity among wild and domesticate accessions should discriminate between different hypotheses of cereal domestication.

To attempt to introduce a view of crop domestication unbiased by the green revolution type of breeding, the nucleotide variation at 18 loci from 92 domesticate einkorn lines compared to 321 lines from wild populations were examined (Kilian et al. 2007b). The data indicate that wild einkorn underwent a natural and strong genetic differentiation, prior domestication, resulting in three distinct *T. boeoticum* races. Only one of those races, indicated as race β , was exploited by humans during domestication. Nucleotide and haplotype diversity in domesticate einkorn was found to be higher than in the β race, supporting the conclusion that einkorn did not undergo a domestication bottleneck. The absence of a domestication bottleneck is in contrast to the conclusions of studies of domestication in intensely bred crop species, where claims for domestication bottlenecks are commonplace. In nature, race β has been sampled only in the “core area” of agricultural development in south-eastern Turkey (Lev-Yadun et al. 2000; Bar-Yosef 2002; Lichter 2007), where the closest wild relatives of einkorn, emmer, barley, rye, chickpea, and lentil still grow (Ladizinsky 1985; Salamini et al. 2002; Ozkan et al. 2005; Abbo et al. 2006). Detailed archaeological reports by Hillman (2000), Willcox (2005) and Weiss et al. (2006) describe how the pre-domestication cultivation of (wild) cereals lasted for centuries in the region, and how it was followed by gradual (Kislev 2002) and multiple (Gebel 2004) appearances of domesticate phenotypes. The genetic and cultural mechanisms underlying the emergence of those phenotypes are remaining questions (Diamond and Belwood 2003).

If geographically distinct domestication events each entailed a random sampling from local genotypes, and if local populations can be identified based on molecular markers, domesticate lines should trace to different localities across the range of the wild progenitor (Jones 2004). This is not observed for einkorn: race β described in Kilian et al. (2007b) appears to be the sister to domesticate einkorn in the absence of an evident reduction of genetic variation. This can be accommodated by a domestication model designated as “dispersed-specific” (Fig. 3.1). In essence, in this scenario a sedentary natufian society (Bar-Yosef 2002) first harvested, and then cultivated the wild β race of *T. boeoticum*, which was distributed only in few locations of the core area. In a later phase of agricultural expansion, the β race was transferred to other locations, possibly already in a state of nascent domestication.

Transport could have involved migrating farmers (Nadel 2002; Renfrew 2002) or exchange of seeds for other goods, as not all soils of the Fertile Crescent were adapted to cereal cultivation (Willcox 2005). Current evidence shows that wild barley seeds were already harvested at the Ohalo camp (in the Palestine corridor) 21,000 years ago (Nadel 2002), indicating that for a long time wild cereals were part of culture and economy of these populations. It follows that particularly hulled forms of barley (Nesbitt and Samuel 1996), were produced and moved across the Fertile Crescent (Willcox 2005).

Concerning einkorn, in several areas variants of the β race emerged having common domesticated traits. These domestication events occurred at several places within the “core area” (These domestication events were connected with

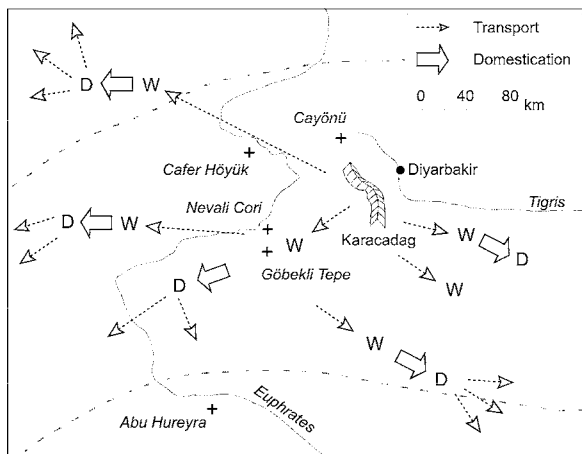


Fig. 3.1 The dispersed-specific model for einkorn domestication. Selected archaeological sites are indicated in *italics*. The Fertile Crescent is indicated with a dotted line. W-wild einkorn (β race); D-domesticated einkorn. In essence, the model assumes pre-adaptation of wild race β for cultivation. This race β was spread to several localities of the Fertile Crescent, before being domesticated at several sites. This figure has been originally published in Kilian et al. 2007b (published with permission from Oxford University Press)

each other due to the same *b* (Beta-symbol font) race seed material used and due to the same human culture) and a genetic bottleneck would have occurred at each domesticating human settlement. However, domestication events at numerous villages would have allowed the newly domesticated lines to integrate a full arsenal of wild haplotypes: in essence, many domestication bottlenecks starting from the same genetic pool of the β race would have resulted in no domestication bottleneck. In this scenario, the specificities still maintained at a molecular level of the β race make it possible to assign the adoption of the β race to the Karacadag core area, while the multiple extraction of domesticates from the β population has preserved a large genetic variation in the domesticated lines.

This hypothesis accounts for our molecular data and accommodates the results of archaeological excavations: tools for grinding seeds are present in the majority of Fertile Crescent sites well before the large seed remains of domesticated einkorn wheat (Bar-Yosef 2002), supporting the view that humans in the region were familiar with the harvest of wild seeds both in natural habitats and in cultivated fields (Weiss et al. 2006; Lichter 2007).

Einkorn is cultivated today on a very small scale as feed for poultry and swine in some mountainous villages in Italy, Spain and Turkey (Nesbitt and Samuel 1996; Perrino et al. 1996). This wheat has been re-discovered as a source of genetic variation for wheat breeding and it is to some extent used by the food industry in Europe (Ozkan et al. 2007).

The second wild diploid *Triticum* species, *T. urartu* (A^uA^u), occurs on basaltic rocks in some parts in the Fertile Crescent (Zohary and Hopf 2000). The species was never domesticated, but played a critical role in wheat

evolution. *T. urartu*, in fact, donated the A genome to all tetraploid and hexaploid wheats (Dvorak et al. 1993).

3.2.1.2 Tetraploid Wheats

Two wild tetraploid wheat species are known, *T. dicoccoides* and *T. araraticum*. They are similar in morphology, but different in their genomic constitution: *T. dicoccoides* has the genomic formula A^uA^uBB and *T. araraticum* A^uA^uGG (Zohary and Hopf 2000).

T. dicoccoides, or wild emmer, has a more restricted distribution range today than wild einkorn and, in nature, it grows especially in the western and central parts of the Fertile Crescent. Wild emmer wheat was for the first time discovered in nature by Aaron Aaronsohn (Aaronsohn and Schweinfurth 1906) and its domesticated form is known as *T. dicoccum* (emmer, A^uA^uBB). The wheat was probably domesticated in South East Turkey (Ozkan et al. 2002, 2005; Mori et al. 2003; Luo et al. 2007). A reconsideration of the domestication geography of tetraploid wheats has been considered by Ozkan et al. (2005) and by Luo et al. (2007). Phylogenetic analysis indicate that two different races of *T. dicoccoides* exist, the western one, colonizing Israel, Syria, Lebanon and Jordan, and the central-eastern one, which has been frequently sampled in Turkey and rarely in Iraq and Iran. It is the central-eastern race that has played the role of the progenitor of the domesticated germplasm. This is supported by the results from the collections of Ozkan et al. (2002), Mori et al. (2003) and Luo et al. (2007), which indicate that the Turkish Karacadag population, has a tree topology consistent with that of the progenitor of domesticated genotypes. The Turkish Kartal population described by Mori et al. (2003) belongs genetically to the central-eastern race but at the nuclear DNA level is less related to the domesticated gene pool. An agreement between published works on tetraploid wheat domestication is evident. A disagreement is nevertheless appearing at the local geographical scale: the chloroplast DNA data indicate the Kartal mountains at the border of the “core area” of plant domestication in the Fertile Crescent (Lev-Yadun et al. 2000), while AFLP fingerprinting points to the Karacadag range as the putative site of tetraploid wheat domestication. From this area, emmer expanded across Asia, Europe and Africa (Kong et al. 2004; Dubcovsky and Dvorak 2007). Southwestern expansion of domesticated emmer generated sympatry with the southern populations of *T. dicoccoides* and the rise of a secondary diversity center (Kong et al. 2004; Luo et al. 2007). This was followed by the subdivision of domesticated emmer into northern and southern subpopulations. Northeast expansion allowed meeting the distribution of *Ae. tauschii* and, thus, the emergence of hexaploid *T. aestivum*. In fact, genetic evidence indicates that the synthesis of hexaploid wheat took place within the corridor from Armenia to the southwestern coast of the Caspian Sea (Gu et al. 2003). Based on the D genome diversity, the synthesis of the hexaploid wheat has been estimated to have occurred at least twice (Dvorak et al. 1998b; Giles and Brown 2006).

Based on a direct estimation of mutation rate for microsatellite loci and re-sequenced candidate loci, Thuillet et al. (2002; 2005) and Haudry et al. (2007) have discussed the occurrence of bottlenecks during tetraploid wheat domestication and breeding. A continuous decrease of effective population sizes is reported, indicating the action of severe bottlenecks, associated in particular to breeding. However, Thuillet et al. (2005) reported that the bottleneck of domestication was relatively low which in terms of Nei's heterozygosity correspond to the presence of 95% of the diversity in domesticated *T. dicoccum* compared to wild *T. dicoccoides* (but 63.2% of effective population size has been lost from wild to domesticated emmer wheat). This situation is remarkably similar to the one reported for einkorn by Kilian et al. (2007b). On the other hand, important losses of nucleotide diversity are reported at 21 loci from the comparisons of domesticate lines of *T. dicoccum* and *T. durum* with the wild *T. dicoccoides* (Haudry et al. 2007). However, in this experiment it is difficult to separate recent bottlenecks from the loss of diversity due only to domestication.

Several cultivated tetraploid A^uA^uBB wheats were derived later from the domesticated emmer: *T. carthlicum* (Persian wheat), *T. polonicum* (Polish wheat), *T. ispahanicum*, *T. turanicum* (Khurasan wheat) and *T. turgidum* (English or pollard wheat). *Triticum dicoccum* was the favored crop for bread-making in ancient Egypt. Like einkorn, emmer wheat cultivation has declined today and it can be found only in some traditional farming communities mainly in Russia and Ethiopia. Somewhat later, *T. durum* (macaroni or hard wheat) originated also from *T. dicoccum* (Damania 1998), possibly independently (Salamini et al. 2002; Ozkan et al. 2005), but different opinions exist on this point (Haudry et al. 2007). This naked wheat is widely cultivated today for pasta production.

In the eastern part of the Fertile Crescent, the wild tetraploid wheat *T. araraticum* (Araratian or Armenian wild emmer) substitutes *T. dicoccoides* (Johnson 1975; Zohary and Hopf 2000). While *T. dicoccoides* crosses easily with cultivated tetraploid wheats, *T. araraticum* does not, most probably due to relevant differences in the genome, like the existence of several translocations between B and G chromosomes (Feldman 1966). *Triticum araraticum* was also domesticated but its cultivated form, *T. timopheevii* (A^uA^uGG; Timopheev's wheat), has been found only in West Georgia together with the hexaploid wheat *T. zhukovskyi* (A^mA^mA^uGG; Zhukovsky's wheat) (Dorofeev et al. 1979). It is speculated that when emmer cultivation spread to Transcaucasia, local populations of *T. araraticum* were colonizing as a weed the fields of emmer crops and, by being incorporated into the agricultural cycle of harvest and sowing, became domesticated (Nesbitt and Samuel 1996).

3.2.1.3 Hexaploid Wheats – Bread Wheat

The most economically important wheat is *T. aestivum* or bread wheat (A^uA^uBBDD). Bread wheat is a temperate crop grown from 67° North in Norway, Finland, and Russia to 45° South in Argentina. *Triticum aestivum* comprises a number of free-threshing forms such as *T. compactum* (club wheat),

T. sphaerococcum (Indian dwarf or shot wheat) *T. petropavlovskyi* (rice wheat) and *T. tibetanum* (Tibetan wheat). Other forms are hulled: *T. spelta* (Dinkel or large spelt), *T. macha*, *T. vavilovii* and *T. yunnanense* (Dvorak et al. 1998a). No wild hexaploid wheat has been ever found, only a semi-wild weedy form of hulled and brittle hexaploid wheat, *T. tibetanum*, has been discovered in Tibet (Shao et al. 1983). It is accepted that *T. aestivum* originated from a cross between domesticated hulled tetraploid emmer *T. dicoccum* (or the free-threshing hard wheat *T. durum*, or the free-threshing *T. parvicoccum*) and the goat grass *Aegilops tauschii* (DD) (Kihara 1944; McFadden and Sears 1946; Kerber 1964; Kislev 1980; Dvorak et al. 1998a; Matsuoka and Nasuda 2004). This cross should have taken place after emmer or hard wheat cultivation spread East from the Fertile Crescent into the natural distribution area of *Ae. tauschii*. The cross occurred most probably South or West of the Caspian Sea about 8000 years ago (Nesbitt and Samuel 1996; Salamini et al. 2002; Giles and Brown 2006). *Aegilops tauschii* encompasses several morphological varieties that are roughly grouped into *Ae. tauschii* ssp. *tauschii* and *Ae. tauschii* ssp. *strangulata* (Kihara et al. 1965; Jaaska 1995; Dvorak et al. 1998a). Several studies show that *Ae. tauschii* ssp. *strangulata* provided the wheat D genome (at least two times), but contributions from both subspecies are also discussed (Nishikawa et al. 1980; Jaaska 1981; Dvorak et al. 1998b; Talbert et al. 1998). If only a few *Ae. tauschii* genotypes participated in the origin of *T. aestivum*, this polyploidization should have been accompanied by reduction of diversity (Haudry et al. 2007). However, high mutation rates, together with buffering effects caused by polyploidy should enable hexaploid wheat to enhance diversity (Dubcovsky and Dvorak 2007).

One still unsolved important question is the origin of hulled hexaploid *T. spelta* or spelt (McFadden and Sears 1946; Kuckuck and Schiemann 1957; Kuckuck 1959; Nishikawa et al. 1980; Dvorak et al. 1998a; Salamini et al. 2002; Blatter et al. 2004). Two types of spelts are known: the Asian and European spelts. Whether *T. spelta* origin is monophyletic or polyphyletic is still open to debate (Nesbitt and Samuel 1996; Dvorak and Luo 2001; Blatter et al. 2004). Genetic data suggest that the hulled hexaploid wheats are more primitive than the free-threshing forms. This hypothesis, however, is not supported by archaeological findings because free-threshing forms occurred earlier than the hulled ones. Hulled wheat appeared in central Europe in the Early Bronze Age (summarized in Nesbitt 2002) and 7,000 year old remains are found in Northern Iraq (Kislev 1984). Free-threshing wheat, in turn, has been found in Can Hassan III dating to 8,500 years ago (Hillman 1978) and at Cafer Höyük dating to 8,000–9,000 years ago (summarized in Salamini et al. 2002). The origin of hulled hexaploid wheat remains controversial.

3.2.2 Barley Evolution and Domestication

Hordeum vulgare (barley) was domesticated from its wild progenitor *H. spontaneum*. Like wheats, barley belongs to the oldest and most important crops of

the Fertile Crescent (Takahashi 1955; Jaaska 1998; Zohary and Hopf 2000; Badr et al. 2000; Bothmer von et al. 2003; Pourkheirandish and Komatsuda 2007). Barley varieties are either two-rowed or six-rowed, based on type of ear. The species is more drought tolerant and much more salt tolerant than wheat. The crop was very important in some regions of the Fertile Crescent, the main crop in Mesopotamia and a primary cereal in ancient Egypt (Harlan 1995).

Wild barley grains have been found in several pre-agricultural PPN sites. The earliest evidence is from Ohalo II, located at the shore of the Sea of Galilee, where 21,000 years old wild remains were found in large amounts (Kislev et al. 1992). This supports the conclusion that wild barley has been collected from nature long before domestication. The earliest carbonized remains of domesticated barley are of the two-row type (van Zeist 1970; Hillman et al. 1989), but six-row types appear already at Ain Ghazal around 9000–8500 years ago (Rollefson et al. 1985; Willcox 1998). Domesticated barley later spread with other crops through the Mediterranean to Europe and Africa, and eastwards through Iran and Afghanistan into India and China.

Wild barley *H. spontaneum* has a wider distribution than any wild wheat. It is present all over the Fertile Crescent because the species is a colonizer of disturbed agricultural habitats. The species occurs in the eastern Mediterranean, western Asia and reaches Turkmenia and Afghanistan in the east (Harlan and Zohary 1966). A few wild barley populations are also found in secondary habitats such as Morocco and Abyssinia.

Considerable studies have been invested to study barley diversity and to identify the region of barley domestication. Badr et al. (2000) reported originally the monophyletic nature of barley domestication based on allelic frequencies at 400 AFLP polymorphic loci studied in 317 wild and 57 domesticated lines. The wild populations from Israel-Jordan were more similar than any others to the domesticated gene pool. The results supported the hypothesis that the Israel-Jordan area was the region in which barley was brought into culture. Moreover, the diagnostic allele I of the homeobox gene *BKn-3* [*Knotted-1*-like-homeobox (*Knox*) gene class], rarely but almost exclusively found in Israeli *H. spontaneum*, was pervasive in western landraces and modern cultivated varieties. In landraces from the Himalayas and India, the *BKn-3* allele IIIa prevails, indicating that an allelic substitution has taken place during the migration of barley from the Fertile Crescent to South Asia. Thus, the Himalayas can be considered a region of domesticated barley diversification. Other reports point to further domestication sites and to different origins (Schieman 1939; Åberg 1940; Bekele 1983; Molina-Cano et al. 1987; Zohary and Hopf 2000; Molina-Cano et al. 2005; Morrell and Clegg 2007; Orabi et al. 2007; Azhaguvel and Komatsuda 2007; Saisho and Purugganan 2007). Allaby and Brown (2003, 2004) questioned the use of AFLP markers in phylogenetic studies addressing crop domestication. Subsequently, Salamini et al. (2004), cited several dozens of papers that correctly addressed domestication issues based on AFLP markers.

Studies based on molecular markers comparing wild to domesticated barley, have shown that a large amount of nucleotide diversity has been lost in current domesticated varieties (Russell et al. 2004; Caldwell et al. 2006; Kilian et al. 2006; Morrell and Clegg 2007). Kilian et al. (2006) determined for a representative sample of 20 domesticated barley (*H. vulgare*) lines and 25 wild *H. spontaneum* lines, the haplotypes at seven loci—*Adh2*, *Adh3*, *Amy1*, *Dhn9*, *GAPDH*, *PEPC* and *WAXY*. The number of haplotypes, average nucleotide diversity, π and Watterson's theta at silent sites was reduced in domesticated lines. Two loci, *Amy1* and *PEPC*, were monomorphic in domesticated lines; *Amy1* and *GAPDH* produced significant values of Tajima's D when all domesticated and wild lines were considered. At *GAPDH*, π was slightly higher in domesticated than wild forms, due to divergent high-frequency haplotypes; for the remaining six loci, 87% of nucleotide diversity has been lost in the domesticated forms. Bottlenecks acting on neutrally evolving loci either during the domestication process, during subsequent breeding, or both, are sufficient to account for reduced diversity and the results of Tajima's test, without the need to evoke selection at these loci. The domesticated varieties considered, although all sampled among those currently cultivated in Turkey, were shown to share the same molecular and morphological variability as larger samples of barley genotypes, indicating that the conclusions reported are of general value for this crop.

Recent data have agreed with the conclusion that two-row and six-row genotypes may have different, independent origins (Zohary and Hopf 2000; Kilian et al. 2006; Komatsuda et al. 2007). These new findings are nevertheless in agreement with the previously inferred area of barley domestication in the Jordan valley (Badr et al. 2000). The new data open the possibility that barley domestication might have been diphyletic. The diphyletic conclusion previously excluded by Badr et al. (2000), is favored by Molina-Cano et al. (2005); Kolodinska Brantestam et al. (2004); Casas et al. (2005); Tanno and Takeda (2004); Komatsuda et al. (2004); Taketa et al. (2004) and Komatsuda et al. (2004). The particular matter concerning single versus multiple origins of barley is, however, complicated by the fact that (i) multiple independent introgressions of genes from wild relatives to cultivated varieties can mimic multiple domestication events (Badr et al. 2000; Kanazin et al. 2002; Abdel-Ghani et al. 2004); and (ii) splitting of domesticated genotypes in two alternative groups based on two-six-rowed ears, hulled-naked caryopsis, western-eastern varieties, and brittleness of the rachis may have followed, and not be coeval with, the domestication process.

3.2.3 Rye Evolution and Domestication

Less is known about rye domestication compared to wheat and barley. Several major questions are still unanswered. Does rye belong to the Neolithic founder crop package or was wild rye incorporated into the agricultural cycle of wheat and became domesticated?

The genus *Secale* includes annual, perennial, autogamous and allogamous species. Wild, weedy, feral and cultivated taxa are recognized. Taxonomy and phylogenetic relationships in *Secale* are controversial and a matter of discussion for more than seven decades (Vavilov 1917; Stutz 1972; Sencer and Hawkes 1980; Hammer et al. 1987; Frederiksen and Petersen 1997, 1998; Jaaska 1998; Zohary and Hopf 2000; Chikmawati et al. 2005). The outcrossing nature and the interspecific-fertility seem to complicate the situation.

Secale is presently recognized as containing three groups of taxa: (i) the annual outcrossing *S. cereale* complex. This group comprises cultivated spring and winter varieties (*S. cereale*; also including the cultivated autogamous *S. turkestanicum*), non-shattering weeds (*S. cereale*), semi-shattering weeds (*S. afghanicum*, *S. dighoricum*, *S. segetale*) and fully shattering wild taxa (including *S. ancestrale*; the inbreeder *S. vavilovii* and the cleistogamous *S. iranicum*); (ii) the wild annual autogamous *S. sylvestre*; and (iii) the wild perennial outcrossing *S. strictum* (syn. *S. montanum*; including *S. africanum*, *S. anatolicum*, *S. chaldicum*, *S. ciliatoglume*, *S. kuprijanovii*) (Sencer and Hawkes 1980; Hammer et al. 1987; Frederiksen and Petersen 1997; Zohary and Hopf 2000; Chikmawati et al. 2005).

Archaeological studies, particularly those describing the findings at Abu Hureyra, indicate that wild rye was already cultivated in the PPN before its domestication (Hillman 2000). Large rye grains with brittle rachis appear at Abu Hureyra about 12,500 years before present. Domesticated grains with tough rachis are present at Can Hasan III at about 8,600 years ago but disappear at other contemporary Neolithic sites in the Near East (Salamini et al. 2002; Weiss et al. 2006). In the Bronze Age, rye is reported at Alaca Höyük, in Anatolia (Hillman 1978). Rye probably reached Europe through a northern route, and remains are present at European Iron Age sites (Salamini et al. 2002; Weiss et al. 2006).

According to Davies and Hillman (1992), the immediate wild (brittle rachis) ancestors of domestic rye is proposed to be *Secale vavilovii*, which is an inbreeder like wheat and barley (Hammer et al. 1987; Hammer 1990). Thus, if brittle rachis *S. vavilovii* lines were exposed to unconscious selections, it would have been domesticated as rapidly as wheat and barley. Probably, wild rye populations invaded fields of cultivated wheat and barley giving rise to weedy forms varying in their degree of rachis brittleness (Vavilov 1917; Sencer and Hawkes 1980). Cultivated rye was then unconsciously or consciously selected starting from weedy types. Sencer and Hawkes (1980) are in favour of the origin of cultivated rye in East Turkey around the Van Lake and that rye was domesticated as a secondary crop several times independently. From East Turkey, the crop spread into Russia, and then from there into Poland and Germany. A second route from South West Asia towards Europe/ across the Balkan Peninsula is also in discussion (Bushuk 2001). Stutz (1972) proposed, based upon extensive cytological, ecological and morphological studies, that cultivated rye originated from weedy progenies derived from introgression of *S. strictum* (syn. *S. montanum*) into *S. vavilovii* and other rye types. *S. vavilovii*

occurs today in Armenia and East Turkey, particularly on the slopes of Mount Ararat and of the Karacadag mountains (Stutz 1972; Sencer and Hawkes 1980; Zohary and Hopf 2000).

A recent phylogenetic study using 789 polymorphic AFLP loci has provided good resolution on the genetic relationships among rye taxa, but did not support the close relationship of *S. vavilovii* with *S. cereale* (Chikmawati et al. 2005).

3.3 Traits Modified by Domestication

Domesticated cereal crops differ from their wild relatives in several traits, some of them apparently consciously selected by humans. The most important *Triticaceae* traits modified during domestication were the free-threshing state and brittle rachis (Table 3.2). Additional modifications taking place during domestication and subsequent breeding concerned seed size, kernel row type, plant height, grain hardness, tillering, seed dormancy, photoperiod, vernalization and heading date. In addition, the spread of the domesticated cereals out of the Fertile Crescent required the adaptation to new environments supported by newly arisen favourable alleles at critical genetic loci.

3.3.1 Free-Threshing

The early wheat varieties were characterized by hulled seeds that required drying to be liberated from the chaff. When species characterized by a low degree of glume tenacity and by fragile rachis and free-threshing habit were selected by the farmers, harvest of grains became efficient. Free-threshing wheats have thinner glumes and paleas which allow an early release of naked kernels. After threshing, free grains are winnowed and stored ready for milling. Free-threshing varieties, like tetraploid hard wheat (*T. durum*) and hexaploid bread wheat (*T. aestivum*) represent the final steps of wheat domestication (Table 3.2).

Also in barley, the hulled/covered or naked caryopsis is an important agronomic trait due to its direct link to dietary use. A single recessive gene, *nud*, located on chromosome 7HL controls the naked caryopsis character. Molecular analysis of a closely linked marker to the *nud* supports the hypothesis of a monophyletic origin of this mutant (Taketa et al. 2004). High-density mapping led to positional cloning of *nud* (Taketa et al. 2006). This group recently showed that an ethylene response factor (ERF) family transcription factor gene controls the covered/naked phenotype. The *Nud* gene has homology to the *Arabidopsis* *WIN1/SHN1* transcription factor, which may control a lipid biosynthesis pathway (Taketa et al. 2008).

Table 3.2 Grass species, genome compositions and genes controlling domestication traits

Species names ¹	Genome and ploidy	Ear and seed traits ²	No. of loci that support B vs. NB rachis ³	Alleles of loci that affect either glume or glume and ear rachis, in brackets the chromosomes ⁴	References
<i>T. boeoticum</i> (w)	A ^a A ^b	H, B	2	<i>SogA</i> (2S)	Sharma and Waines 1980; Taenzler et al. 2002
<i>T. monococcum</i> (d)	A ^m A ^m	H, NB	2	<i>SogA</i> (2S); <i>Q</i> (5AL)	Sharma and Waines 1980; Taenzler et al. 2002; Faris et al. 2003; Simons et al. 2006
<i>T. urartu</i> (w)	A ^a A ^u	H, B	2		Simons et al. 2006
<i>Ae. speloides</i> (w)	BB	H, B	1	<i>Br-S1</i> (3S); <i>Br1</i> (W-type) or <i>Br2</i> (B-type)	Li and Gill 2006
<i>Ae. tauschii</i> (w)	DD	H, B	1	<i>Br2</i> (3DL); <i>TgD</i> (2S)	Kerber and Rowland 1974; Villareal et al. 1996; Jantasuriyarat et al. 2004;
					Watanabe et al. 2005a, b, 2006b;
					Li and Gill 2006
<i>T. dicoccoides</i> (w)	A ^a A ^u BB	H, B	2, polygenic	<i>Tg2B</i> (2S); <i>qA</i> (5L); <i>Qft5A</i> = <i>QA</i> (5S); <i>Qft6A</i> (6); <i>Br-A1</i> (3AS); <i>Br-B1</i> (3BS)	Sears 1954; Schiemann and Staudt 1958; Sears 1976; Muramatsu 1985;
					Muramatsu 1986; Chen et al. 1998; Simonetti et al. 1999; Watanabe and Ikebata 2000; Watanabe et al. 2002; Nalam et al. 2006; Simons et al. 2006
<i>T. dicoccum</i> (d)	A ^a A ^u BB	H, NB	2	<i>Tg2B</i> (2S); <i>qA</i> (5L)	Muramatsu 1985; Muramatsu 1986; Nalam et al. 2006; Simons et al. 2006
<i>T. durum</i> (d)	A ^a A ^u BB	FT, NB	polygenic	<i>tg2B</i> (2S); <i>QA</i> (5L); <i>qft5A</i> (5S); <i>qft6A</i> (6)	McKey 1966; Kerber and Rowland 1974; Sears 1976; Muramatsu 1985;
					Muramatsu 1986; Simonetti et al. 1999; Watanabe and Ikebata 2000;
<i>T. araraticum</i> (w)	A ^a A ^u BB	H, B	1	<i>Br1</i> (3AS)	Watanabe 2005; Watanabe et al. 2005a; Simons et al. 2006; Nalam et al. 2007
<i>T. tinopheevii</i> (d)	A ^a A ^u BB	H, NB			Li and Gill 2006
<i>T. spelta</i> (d)	A ^a A ^u BBDD	H, NB	2	<i>Tg2B</i> <i>Tg2D</i> (2S); <i>qA</i> (5L); <i>Qft5A</i> (5S); <i>Qft6A</i> (6)	Li and Gill 2006
					Sears 1976; Ternowskaya and Zhirov 1993; Bömer and Worland 1996;
					Cao et al. 1997; Iqbal et al. 2000; Luo et al. 2000; Kato et al. 2003; Li and Gill 2006;
					Onishi et al. 2006; Simons et al. 2006

Table 3.2 (continued)

Species names ¹	Genome and ploidy	Ear and seed traits ²	No. of loci that support B vs. NB rachis ³	Alleles of loci that affect either glume or glume and ear rachis, in brackets the chromosomes ⁴	References
<i>T. aestivum</i> (d)	A ^u A ^u BBDD	FT, NB	2	<i>tg2b/g2b</i> ; <i>QAL</i> ⁸ ; <i>qf15A</i> ⁸ ; <i>qf16A</i> ⁸ ; <i>br-A1</i> (3AS); <i>br-B1</i> (3BS); <i>br-D1</i> (3DS)	MacKey 1954; Sears 1954; Kerber and Rowland 1974; Muramatsu 1986; Faris and Gill 2002; Watanabe et al. 2002; Kato et al. 2003; Jantasuriyarat et al. 2004; Watanabe et al. 2005a; Li and Gill 2006; Simons et al. 2006; Watanabe et al. 2006a; Nalam et al. 2007 Takahashi 1955; Takahashi 1972; Kandemir et al. 2004; Komatsuda et al. 2004; Azhaguvel et al. 2006; Azhaguvel and Komatsuda 2007; Komatsuda et al. 2007; Pourkheirandish and Komatsuda 2007 Takahashi 1955, 1972; Sogaard and Wettstein-Knowles von 1987; Kandemir et al. 2004; Komatsuda et al. 2004; Azhaguvel et al. 2006; Azhaguvel and Komatsuda 2007; Komatsuda et al. 2007; Pourkheirandish and Komatsuda 2007 Kranz 1963; Frederiksen and Petersen 1998 Kranz 1963; Frederiksen and Petersen 1998
<i>H. spontaneum</i> (w)	HH	H, B	2	<i>Btr1</i> (3H); <i>Btr2</i> (3H)	
<i>H. vulgare</i> (d)	HH	H ⁵ , NB	2	<i>btr1</i> (3H); <i>btr2</i> (3H) ⁹	
<i>S. vavilovii</i> (w)	RR (SS)	B	1		
<i>S. cereale</i> (d)	RR (SS)	NB	1		

¹ See text for the wheat classification system used in this chapter. w-wild; d-domesticated; ² H-hulled; FT-free-threshing; B-brittle rachis; NB-non-brittle rachis; ³ Genes that affect rachis but not glume traits; ⁴ Subscribis indicate genomes; ⁵ The trait hulled seeds (H) is under the control of the single gene N (recessive n, naked seeds) – Takahashi (1955); ⁶ Designated as *q2* in Simonetti et al. (1999); ⁷ Allele inferred from genotype of wild emmer; ⁸ Allele inferred from genotype of hard wheat; ⁹ Most occidental cultivars are of genotype *btr1Btr2*, while most oriental ones are *Btr1btr2* (Takahashi 1955).

Major and minor mutations have been considered to explain the evolution of the free-threshing habit in wheat (McKey 1966; Jantasuriyarat et al. 2004). The major component is the *Q* gene or gene complex, located on chromosome 5AL. In *T. aestivum*, the *Q* allele supports the formation of square-headed ears with good threshability, besides inducing softening of the glumes, reduction of ear length, more spikelets per ear, and toughness of the rachis (Sears 1954; Snape et al. 1985; Kato et al. 1998; 2003). Disruption of the *Q* gene generates a *q* mutant phenotype, known as speltoid type because *q* mutants have tenacious glumes similar to that of spelt (*T. spelta*; *qq* genotype). Bread wheat lines harbouring both *Qq* alleles have intermediate phenotypes. Muramatsu (1963) also showed that the *q* allele is active by creating genotypes with 1–5 doses of either *Q* or *q* alleles. He showed that a square-headed hexaploid ear derives from either two doses of *Q* or five doses of *q*.

In hexaploid wheats, the polygenic component controlling free-threshing is scattered throughout all three genomes. In tetraploid wheats, QTL (quantitative trait locus) studies identified a total of four putative loci (Simonetti et al. 1999), located on chromosomes 2B, 5A, and 6A. Two of these QTLs correspond, in position, to the already mentioned gene *Q* and to the genetic locus *Tg* (tenacious glumes). A recent mapping effort led to the identification of two QTL affecting both glume adherence and threshability (Nalam et al. 2006), suggesting that threshability is a function of glume adherence (Nalam et al. 2007).

The *Q* gene was cloned and shown to have sequence similarity to the *Arabidopsis APETALA2* gene (Faris and Gill 2002; Faris et al. 2003). Further studies confirmed the association (Simons et al. 2006) and demonstrated that ectopic expression of *Q* in transgenic plants mimicked dosage and pleiotropic effects of *Q*. Increased transcription of *Q* was associated with spike compactness and reduced plant height. Previous research suggested that *Q* might have arisen from a duplication of *q* (Kuckuck 1959). However, Simons et al. (2006) contradict this hypothesis, showing that most probably *Q* arose through a gain-of-function mutation.

A second major gene, *Tg*, controls threshability and is associated with the wheat D genome. *Tg* controls the speltoid phenotype and inhibits the expression of *Q*. The suppression of the free-threshing character was thought to be due to a partially dominant *Tg* allele on chromosome 2D, contributed from *Ae. tauschii* and thus leading to tenacious glumes. The conclusion is that free-threshing hexaploids have the genotype *tggt*, *QQ* (Kerber and Rowland 1974; Villareal et al. 1996).

Rachis fragility and square-headed ears in free-threshing hexaploid wheats can assort independently. Moreover, *q* and *Q* interact with genes on other chromosomes that also govern glume tenacity - rachis fragility (Cao et al. 1997; Luo et al. 2000). Another gene, located 32 cM from *Tg*, on chromosome 2D, belongs to the system controlling and regulating the same trait.

Further major genes responsible for compact spikes are known: *C* on the long arm of chromosome 2D in *T. compactum*, *C2* in *T. antiquorum* and *T. sphaerococcum* (Goncharov et al. 2008) and *S-D1* on the long arm of chromosome 3D (Kerber and Dyck 1969; Rao 1972; Kerber and Rowland 1974; Rao 1977).

3.3.2 *Brittle-Rachis*

The modification of the brittle rachis trait has been critical for the origin of agriculture and sedentary societies. In nature, the spikelets of the wild ears fall apart at ripening through fragmentation of the rachis (by shattering or disarticulation). This mechanism is necessary for seed dispersal and self-planting. In a tough, non-brittle rachis the formation of fracture zones at the rachis is suppressed until mature spikes are harvested by man. It is thought that the spikes of non-brittle mutated plants were consciously selected by early farmers and that their frequency increased constantly in cultivated fields.

In wheats, the brittle rachis character maps to the homeologous group 3 chromosomes (Watanabe et al. 2002; Salamini et al. 2002; Watanabe 2005; Li and Gill 2006). In einkorn, this trait is under the control of two genes that segregate 15 brittle to 1 tough rachis in the F₂ progeny of wild × domesticated crosses (Sharma and Waines 1980). Cao et al. (1997) identified a single dominant gene, *Br-A1* (*Br₁*, brittle rachis), responsible for rachis fragility in a feral form of *T. aestivum* from Tibet. The gene was later localized on chromosome 3DS (Chen et al. 1998), as supported by studies of a cross of *T. dicoccoides* × *T. aestivum* (Rong et al. 2000). Further dominant genes are *Br₂* and *Br₃* on chromosomes 3A and 3B, respectively (Cao et al. 1997; Chen et al. 1998; Watanabe and Ikebata 2000).

The studies reported point to (i) multiple genetic pathways controlling the trait(s); and (ii) different genetic origins of loci controlling shattering in polyploids (Salamini et al. 2002). These considerations, combined with the mapping of QTLs for shattering, allow the study of microsyntenous relationships of these traits in the *Triticeae* and other grasses.

Brittle rachis in wild barley (*H. spontaneum*) is controlled by two dominant complementary genes, *Btr1* and *Btr2*. These two genes are tightly linked on chromosome 3H (Azhaguvel et al. 2006). During barley domestication the recessive *btr1* and *btr2* alleles were selected, probably independently from each other. The *btr1* allele is present in most western cultivars (*btr1Btr2*) whereas *btr2* occurs in most eastern cultivars (*Btr1btr2*) (Azhaguvel and Komatsuda 2007). The non-brittle rachis phenotype is probably due to a loss of function of the two genes (Komatsuda et al. 2004). The *btr1* (non-brittle rachis) gene has recently been fine mapped (Azhaguvel et al. 2006).

3.3.3 *Seed Size and Grain Yield*

A transition from small-seeded wild plants with natural seed dispersal to larger-seeded non-shattering plants is evident. Seed size, and thus grain yield, was positively selected in domesticated cereals. The genetic control of seed size in domesticated versus wild tetraploid wheats was studied by generating *T. dicoccoides* substitution lines in *T. durum* (Elias et al. 1996). Kernel size is under a

complex polygenic control, and genes with alleles contributing to an increase and to a decrease in kernel size have been mapped to chromosomes 1A, 2A, 3A, 4A, 7A, 5B, and 7B. In a recent experiment (Peng et al. 2003), eight QTLs for grain weight were mapped, each orthologous to rice counterparts. QTLs responsible for *T. aestivum* yield traits were also mapped to chromosomes 3A, 4A, and 5A (Shah et al. 1999; Campbell et al. 2003; Araki et al. 1999; Kato et al. 2000). In a recent study (Brescaghello and Sorrells 2006), the association of 36 unlinked simple-sequence repeat (SSR) markers with kernel size and milling quality was analyzed in a collection of modern cultivars of soft winter wheat and QTLs for kernel size were detected on chromosomes 2D and 5A/5B.

3.3.4 Kernel Rows in the Ear

One additional conspicuous case of selection in barley domestication was the six-rowed spike responsible for significant yield increase. Archaeological and genetic evidence shows that six-rowed barley was derived from two-rowed barley, probably very early in PPN. At least five independent loci control the six-rowed spike phenotype in barley (summarized in Pourkheirandish and Komatsuda 2007). The most important one, the recessive *vsr1* gene, present in all six-rowed cultivars, was recently cloned by Komatsuda et al. (2007) and is located on chromosome 2HL. Wild barleys and cultivated two-rowed barleys have a dominant *VRS1* (six rowed spike) gene (summarized in Pourkheirandish and Komatsuda 2007). Closely linked markers to *vsr1* indicate that six-rowed barley originated more than once during domestication (Tanno et al. 2002).

3.3.5 Plant Height

The “green revolution” in grasses was achieved by reducing the lodging susceptibility and increasing the grain yield (reviewed in Hedden 2003). In rice, this was achieved by introducing mutations in genes controlling a key-step in gibberellin acid (GA) biosynthesis (Salamini 2003; Pozzi et al. 2004). Modern wheats are short because they respond abnormally to gibberellin. The *Rht-1* (reduced height-1) gene in wheat encodes a repressor of GA signalling orthologous to *Arabidopsis GAI* (gibberellic acid “insensitive”), maize *dwarf8* (*d8*) and barley *Slender1* (*Sh1*) (Peng et al. 1999; Chandler et al. 2002; reviewed in Hedden 2003; Eastmond and Jones 2005). Interestingly, *d8* of maize is a target of selection and adaptation to flowering time (Thornberry et al. 2001). Pleiotropic effects are not surprising for genes controlling hormone action and may be a common occurrence for the traits targeted by domestication and breeding (Cai and Morishima 2002; Salamini et al. 2002).

Rht-B1b and *Rht-D1b* genes on wheat chromosomes 4B and 4D are semi-dominant, mutant alleles of the *Rht-1* gene conferring dwarfism (reviewed in Hedden 2003). They are present today in most wheat varieties, have been

transferred from the Japanese variety Norin 10 into CIMMYT germplasm, and then used for world-wide breeding programs. *Rht-8*, also capable of reducing plant height by 10 cm, is derived from the Japanese variety Akakomuki. It is likely that the Korean peninsula was the centre of origin of these useful mutants (Cho et al. 1993).

Other height reducing genes are known and *Rht* genes may reach the number of 20. In addition, genes were identified that reduce plant height without affecting early growth, or coleoptile length and vigour (Ellis et al. 2005). In a recent study, these genes were mapped to different wheat chromosomes, thus widening their exploitation in plant breeding (Ellis et al. 2005).

3.3.6 Grain Hardness

Grain hardness or texture is an important quality trait of the grain. The trait controls mainly the susceptibility to damage during milling and the amount of water uptake during baking. Compared to *T. aestivum*, *T. durum* (hard wheat) has hard endosperm and is therefore mainly used for pasta production.

The locus influencing hardness (i.e. conferring the “soft” phenotype) named *Ha* (hardness) was mapped to the short arm of chromosome 5D (Sourdille et al. 1996) and encodes mainly friabilins included in the prolamin superfamily: puroindoline a (*Pina*), puroindoline b (*Pinb*), and grain softness protein (*Gsp-1*) (Gautier et al. 1994; Rahman et al. 1994; Kan et al. 2006). These three genes are tightly linked (Sourdille et al. 1996; Giroux and Morris 1998; Giroux et al. 2000; Chantret et al. 2004) and their physical order (*Pinb-Pina-Gsp-1*) was investigated in bacterial artificial chromosome (BAC) clones in several *Triticum* and *Aegilops* species (Tranquilli et al. 1999; Turnbull et al. 2003; Chantret et al. 2004, 2005; Li et al. 2008). The molecular basis and the evolutionary events at this locus were investigated. The variation at the *Ha* locus arose from gene loss after polyploidization. Genomic rearrangements, such as transposable element insertions, deletions, duplications and inversions contributed to the differences between species of different ploidy levels (Chantret et al. 2005; Massa and Morris 2006). *Pina* and *Pinb* genes are conserved in all diploid *Triticum*, *Aegilops* and closely related *Pooideae*, but are deleted from the A^u and B chromosomes in tetraploid A^uA^uBB wheats (Gautier et al. 2000; Li et al. 2008). *Pina* and *Pinb* gene loss from the *Ha* locus of A^uA^uBB tetraploids was caused by a large genomic deletion that probably occurred independently in the A^u and B progenitor genomes of the A^uA^uBB tetraploid wheats (Li et al. 2008). However, *Pina* and *Pinb* genes are present on chromosome 5D in hexaploid wheats, because they were introduced by *Ae. tauschii*, at the time of the emergence of the allohexaploid bread wheat. Recently, Li et al. (2008) showed that *Pina* and *Pinb* were eliminated from the G genome, but maintained in the A^u genome of tetraploid A^uA^uGG wheats. This supports the independent polyploidization events leading to A^uA^uBB and A^uA^uGG wheats.

The *Gsp-1* genes are conserved in genomes A^u, B, D and G at all ploidy levels. They constitute a multigene family and may be functionally important, particularly in *T. durum* where they have major roles in plant defence and only a minor influence on grain texture (Gollan et al. 2007)

After the acquisition of *Ha* by hexaploid wheats via the D genome, the spreading of the species to northern cultivation areas forced breeders to select hexaploid hard wheat cultivars, and thus the selection pressure on *Ha* has increased. Hexaploid hard wheats have a mutation in either *Pina* or *Pinb* but not in *Gsp-1* (Giroux and Morris 1998). In addition, other genetic factors unlinked to the *Ha* locus may contribute to modify grain hardness (Perretant et al. 2000). The recent knowledge on puroindolines is summarized in Morris (2002) and Bhavé and Morris (2008a,b).

3.3.7 Tillering

Tillering and branching mutations have been described in barley, where; for example, the recessive *uniculm2* mutation on chromosome 6H fails to produce tillers (Babb and Muehlbauer 2003). The genetic variation for the capacity to tiller was assessed for the wheat gene pool (Atsmon and Jacobs 1977): low tillering genotypes frequently have a unicum phenotype, enlarged spike and modified leaf morphology. Some wild diploid grasses belonging to the *Aegilops* section *Sitopsis* (like *Ae. speltoides*) have an ear morphology mimicking the teosinte (wild teosinte, *Zea mays* ssp. *parviglumis*, is most closely related to domesticate maize, *Z.m.ssp. mays*, and is expected to be the progenitor of maize). It is concluded that current wheats have already undergone changes leading to a maize-cob like morphology. In wheat, a single recessive gene (*tiller inhibition gene*; *tin*) located on chromosome 1AS was found to control tiller number (Spielmeyer and Richards 2004). In *tin* populations, low-tillering lines have a greater harvest index, fewer and sterile tillers, and a larger grain size, demonstrating the agronomic potential of the *tin* gene (Richards 1988; Duggan et al. 2002). Comparative genomics analyses lead to the conclusion that *tin*, rice reduced tillering mutations and the barley *uniculm2* mutant map to non-syntenic chromosomes (Rossini et al. 2006).

Rescently, a tiller inhibition gene *tin3* has been identified in diploid wheat *T. monococcum* by ethyl methanesulphonate (EMS)-based mutagenesis. The tillering mutant produces only one main culm compared to the wild type with many tillers. The monicum phenotype of *tin3* is due to a single recessive mutation. Genetic and molecular mapping located the *tin3* gene on the long arm of chromosome 3A^m (Kuraparthi et al. 2007). In their present paper the same group reports the genomic targeting and mapping of *tin3* gene using wheat ESTs (expressed sequence tags) and synteny to rice. The *tin3* gene was mapped to a 324-kb region of rice chromosome arm 1L (Kuraparthi et al. 2008).

3.3.8 *Reduced Seed Dormancy*

Dormancy (i.e. the delay of seed germination) enables wild plants to survive under unsuitable environmental conditions, such as in winter or during dry seasons. Seeds of wild plants are highly dormant (Takeda 1995), whereas for modern cultivars rapid seed germination is required, for instance for the malting process (summarized in Pourkheirandish and Komatsuda 2007). Seed dormancy is a complex trait in barley due to the number of genes participating, the influence of environmental factors during seed development, seed storage and seed germination, and the different after-ripening requirements of the seeds (Alonso-Blanco et al. 2003). Several dormancy loci were identified by QTL analysis in barley. Two major QTLs, *SD1* and *SD2* (seed dormancy), both located on the long arm of chromosome 5H, show the largest effects (summarized in Pourkheirandish and Komatsuda 2007).

3.3.9 *Control of Flowering Time*

The control of flowering time is central to reproductive success and has a major impact on grain yield in *Triticeae*. Wild progenitors of domesticated cereals are well adapted to the prevailing environmental conditions in the Fertile Crescent. These species use the rainfall in fall to establish their vegetative structures before winter, thereby using vernalization to delay flowering until the winter has passed to prevent frost damage. So, in spring the moisture can be used for grain filling to flower early in response to long days (LDs) ahead of the hot summer (Cockram et al. 2007).

The first cereals domesticated in the Fertile Crescent presumably showed the photoperiodic (response to day length) and vernalization (response to winter temperature) phenotypes of their progenitors. However, during the domestication process and during the spread of agriculture out of the Fertile Crescent novel adaptive traits suited for the new environments were selected. One key event was the selection of spring types that can be sown after winter. These spring types lack vernalization requirement and show different response to LDs. Reduced photoperiod response is important in Europe and North America, where growing seasons are long (Turner et al. 2005).

Wheat, barley and *Arabidopsis* show similar flowering responses to photoperiod and vernalization, in contrast to rice, which flowers in response to short days and shows no response to vernalization.

3.3.10 *Photoperiod*

The major loci affecting the photoperiod response (*Ppd* genes) are in a collinear position on the short arm of the group 2 chromosomes in wheat and barley. In

barley, dominant *Ppd-H1* alleles confer early flowering under LDs, but have no effect under short days (SD). Plants carrying the mutated, recessive *ppd-H1* allele are late-flowering. The *Ppd-H1* is a pseudo-response regulator (PRR) most similar to *PRR7* in *Arabidopsis* and was recently positionally cloned in barley by Turner et al. (2005). These authors provided evidence that a single point mutation within the conserved CCT domain of PRR results in an amino acid change that results in insensitivity under LDs. In wheat, the allelic series of *Ppd* loci has decreasing potency from *Ppd-D1* to *Ppd-B1* to *Ppd* to *A1* (Worland 1996). The *Ppd-H2* gene has been mapped in barley on chromosome 1H (Laurie et al. 1995). The winter allele delays flowering under SD. No equivalent gene has been identified in wheat yet.

Further major photoperiod related genes/gene families appear to be conserved between barley and *Arabidopsis*, involving the *GIGANTEA* (*GI*), *CONSTANS* (*CO*), and *FLOWERING LOCUS T* (*FT*) genes in *Arabidopsis* and their orthologs in barley *HvGI*, *HvCO* and *HvFT* (Griffiths et al. 2003; Dunford et al. 2005; Cockram et al. 2007; Faure et al. 2007). Nevertheless, none of the grass QTLs associated with flowering time cosegregate with orthologous *Arabidopsis* “flowering” genes, i.e. different major determinants of photoperiod have been selected in the *Triticeae* (Börner et al. 1998; reviewed in Griffiths et al. 2003).

At the moment, *Ppd-H1* is the only locus conferring natural genetic variation in photoperiod response in barley. Jones et al. (unpublished; cited in Cockram et al. 2007) studied molecular diversity at the *Ppd-H1* locus in a comprehensive barley collection. The mutated non-responsive, late-flowering allele *ppd-H1* was not present in wild barley, suggesting that the mutation occurred after domestication during the spread of agriculture (Cockram et al. 2007).

3.3.11 Vernalization

In the *Triticeae*, two major genes, with epistatic interactions, mainly control the vernalization (*VRN*) pathway: *VRN1* and *VRN2* (Yan et al. 2003, 2004a,b; von Zitzewitz et al. 2005; Dubcovsky et al. 2006). The *VRN1* gene on chromosome 5 in wheat and barley is similar to the *Arabidopsis* MADS-box transcription factor *Apetala 1* (*API*), which initiates the transition of the apical meristem from the vegetative to the reproductive state (Yan et al. 2003). Mutations in regulatory regions (promoter and the first intron) are associated with a dominant spring growth habit (*Vrn1*) (Yan et al. 2003, 2004a; Fu et al. 2005; von Zitzewitz et al. 2005). The second vernalization gene, *VRN2* (*ZCCT* in barley) is a zinc finger CCT domain transcription factor that blocks the photoperiod pathway through direct or indirect down-regulation of *VRN1* and *VRN3* (Fu et al. 2005; Karsai et al. 2005; Koti et al. 2006). *VRN2* transcription is repressed by vernalization and by short days. Loss-of function mutations (*vrn2a*), or complete gene deletion (*vrn2b*), are associated with a recessive spring habit that do not require vernalization to flower (Yan et al. 2004b; Dubcovsky et al. 2005).

It was recently shown that a third vernalization locus in barley is located on chromosome 7H, *VRN-H3*, collinear with *VRN-B3* in bread wheat and encoding an ortholog of the *Arabidopsis* floral pathway integrator, *FT1*, and collinear with *OsFT* of rice (Yan et al. 2006; Faure et al. 2007). The dominant, early flowering, wheat *Vrn3* allele is associated with the insertion of a retroelement in the promoter, while in barley, a mutation in the first intron is associated with the early allele (Yan et al. 2006). This study provides evidence that differences in flowering time are associated with *FT1* allelic variation. The link between photoperiod and vernalization is evident, even at the molecular level, if one considers that *VRN2* is expressed only when photoperiod-responsive plants are grown under long-day photoperiod (Trevaskis et al. 2006). There is probably a second *VRN1* repressor and a candidate may be *VRT2* (*vegetative to reproductive transition-1*), a MADS-box gene (Danyluk et al. 2003; Kane et al. 2005; Szűcs et al. 2006, 2007).

3.3.12 Heading Time

Additional photoperiod and temperature sensitivity loci, responsible for flowering time fine-tuning, such as *earliness per se* on chromosome 1A^{mL} in *T. monococcum*, are reviewed in Cockram et al. (2007). These are responsible for the fine-tuning of wheat flowering time. Bullrich et al. (2002) have mapped to chromosome 1A^{mL} the “*earliness per se*” gene *Eps-Aml* in *T. monococcum*.

QTL mapping for loci controlling flowering time has been carried out by Peng et al. (2003) in a cross between the wild *T. dicoccoides* and the domesticated *T. durum*. The authors identified up to 76 QTL effects. Four flowering time QTLs mapped on chromosomes 2A, 4B, 5A and 6B. The wild parent was sensitive to day length, and flowered later compared to the cultivar. The authors conclude that the wild allele for the QTL on 5A is responsible for late flowering of *T. dicoccoides*, whereas the wild alleles on 2A, 4B, and 6B accelerate flowering date.

3.3.13 Conclusions and Final Considerations

Archaeological and genetic evidence indicate that western agriculture began in the Fertile Crescent about 12,000 years ago (Zohary and Hopf 2000; Salamini et al. 2002). The present view is that the process of crop domestication was slow, spanned several centuries and entailed repeated domestication events (Tanno and Willcox 2006). Local wild populations were domesticated in a core area of the Fertile Crescent and were then gradually dispersed throughout the region (Abbo et al. 2006). All current domestication models predict a reduction of genetic diversity in domesticate forms compared to their wild progenitors (Doebley et al. 2006), but evidence from studies sampling large numbers of

plants and loci are lacking. Recently Kilian et al. (2007b) provided evidence that at least einkorn underwent no reduction of diversity during its domestication, based upon identifying a wild einkorn race representing the wild progenitor among the whole wild gene pool of *T. boeoticum*.

The keys to obtain deeper insights to plant domestication using molecular biology are (i) a comprehensive germplasm collection covering the whole distribution area for each species; (ii) the comparison of many wild and domesticated accessions for each species; (iii) the identification of the wild progenitor in the wild gene pool and its comparison with domesticate descendants; (iv) the use of new molecular fingerprinting techniques at many loci and the access to new generation high throughput sequencing technologies (Goldberg et al. 2006; Wicker et al. 2006); and (v) improvement of analytical methods capable of treating domestication issues based on mathematical and statistical models (Pluzhnikov and Donnelly 1996; Thuillet et al. 2005; Haudry et al. 2007).

New genomic resources for future plant breeding have to be developed and agronomically-important genes are being isolated. International consortia, such as the International Triticeae Mapping Initiative (ITMI), the International Wheat Genome Sequencing Consortium (IWGSC), the International Barley Sequencing Consortium (IBSC), and ERA-PG (European Research Area Networks - Plant Genomics) founded consortia like (EXBARDIV) (Genomics-Assisted Analysis and Exploitation of Barley Diversity) will lead to accelerated gene discovery and will shed new light on mechanisms that have shaped the wheat, barley and rye genomes during their evolution and domestication.

Ongoing archaeological excavations, like at Göbekli Tepe in South East Turkey (Schmidt 2001, 2006), will provide new information on cultural/religious backgrounds at the time of the rise of agriculture, thereby producing insights why humans choose those directions that led to our present-day culture.

Acknowledgments We thank Sigi Effgen, Isabell Fuchs, Jutta Schütze, Charlotte Bulich, Marianne Haberscheid for excellent technical assistance and Margit Pasemann, Birgit Thron, Marianne Limpert, Elke Bohlscheid, Katuscia Ceron for administration support during the last years. We are grateful to the MPIZ sequence facilities (ADIS) headed by Bernd Weisshaar. We thank Bill Martin, Maarten Koornneef, George Coupland, Moshe Feldman, Andrea Brandolini, Klaus Schmidt (DAI) and Andreas Graner for valuable suggestions. This research was supported by the Deutsche Forschungsgemeinschaft SPP 1127.

References

- Aaronsohn, A. and Schweinfurth, G. (1906) Die Auffindung des wilden Emmers (*Triticum dicoccum*) in Nordpalästina. *Altneuland Monatsschrift für die Wirtschaft. Erschliessung Palästinas* 7–8, 213–220.
- Abbo, S., Gopher, A., Peleg, Z., Saranga, Y., Fahima, T., Salamini, F. and Lev-Yadun, S. (2006) The ripples of “The Big (agricultural) Bang”: The spread of early wheat cultivation. *Genome* 49, 861–863.

- Abdel-Ghani, A.H., Parzies, H.K., Omary, A. and Geiger, H.H. (2004) Estimating the out-crossing rate of barley landraces and wild barley populations collected from ecologically different regions of Jordan. *Theor. Appl. Genet.* 109, 588–595.
- Åberg, E. (1940) The taxonomy and phylogeny of *Hordeum* L. sect. *Critesion* Ands. with special reference to Tibetan barleys. *Symb. Bot. Upsaliensis* 2, 1–156.
- Allaby, R.G. and Brown, T.A. (2003) AFLP data and the origins of domesticated crops. *Genome* 46, 448–453.
- Allaby, R.G. and Brwon, T.A. (2004) Reply to the comment by Salamini et al. on “AFLP data and the origins of domesticated crops”. *Genome* 47, 621–622.
- Alonso-Blanco, C., Bentsink, L., Hanhart, C.J., Blankestijn-de Vries, H. and Koornneef, M. (2003) Analysis of natural variation at seed dormancy loci of *Arabidopsis thaliana*. *Genetics* 164, 711–729.
- Ammerman, A.J. and Cavalli-Sforza, L.L. (1984) *The neolithic transition and the genetics of populations in Europe*. Princeton, Princeton University Press.
- Araki, E., Miura, H. and Sawada, S. (1999) Identification of genetic loci affecting amylose content and agronomic traits on chromosome 4A of wheat. *Theor. Appl. Genet.* 98, 977–984.
- Atsmon, D. and Jacobs, E. (1977) A newly bred ‘Gigas’ form of bread wheat (*Triticum aestivum* L.): morphological features and thermo-photoperiodic responses. *Crop Sci.* 17, 31–35.
- Azhaguvel, P., Vidya-Saraswathi, D. and Komatsuda, T. (2006) High-resolution linkage mapping for the non-brittle rachis locus *btr1* in cultivated × wild barley (*Hordeum vulgare*). *Plant Sci.* 170, 1087–1094.
- Azhaguvel, P. and Komatsuda, T. (2007) A phylogenetic analysis based on nucleotide sequence of a marker linked to the brittle rachis locus indicates a diphyletic origin of barley. *Ann. Bot.* 100, 1009–1015.
- Babb, S. and Muehlbauer, G.J. (2003) Genetic and morphological characterization of the barley *uniculm2* (*cul2*) mutant. *Theor. Appl. Genet.* 106, 846–857.
- Badr, A., Müller, K., Schäfer-Pregl, R., El Rabey, H., Effgen, S., Ibrahim, H.H., Pozzi, C., Rohde, W. and Salamini, F. (2000) On the origin and domestication history of barley (*Hordeum vulgare*). *Mol. Biol. Evol.* 17, 499–510.
- Bar-Yosef, O. (2002) The Natufian culture and the early Neolithic – Social and economic trends. In: P. Bellwood, C. Renfrew (Eds.), *Examining the farming/language dispersal hypothesis*. McDonald Institute for Archaeological Research, Cambridge, pp. 113–126.
- Bekele, E. (1983) A differential rate of regional distribution of barley flavonoid patterns in Ethiopia, and a view on the center of origin of barley. *Hereditas* 98, 269–280.
- Bhave, M. and Morris, C.F. (2008a) Molecular genetics of puroindolines and related genes: regulation of expression, membrane binding properties and applications. *Plant Mol. Biol.* 66, 221–231.
- Bhave, M. and Morris, C.F. (2008b) Molecular genetics of puroindolines and related genes: allelic diversity in wheat and other grasses. *Plant Mol. Biol.* 66, 205–219.
- Blatter, R.H.E., Jacomet, S. and Schlumbaum, A. (2004) About the origin of European spelt (*Triticum spelta* L.): allelic differentiation of the HMW Glutenin B1-1 and A1-2 subunit genes. *Theor. Appl. Genet.* 108, 360–367.
- Börner, A. and Worland, A. (Eds.) (1996) Selected papers from the EWAC-conference: cereal aneuploids for genetic analysis and molecular techniques. *Euphytica* 89, 1–157.
- Börner, A., Korzun, V. and Worland, A.J. (1998) Comparative genetic mapping of loci affecting plant height and development in cereals. *Euphytica* 100, 245–248.
- Bothmer von, R., Hintum van, T., Knüpfper, H. and Sato, K. (Eds.), (2003) *Diversity in barley (Hordeum vulgare)*. Elsevier, Amsterdam.
- Braidwood, R.J. and Braidwood, L. (1950) Jarmo: A village of early farmers in Iraq. *Antiquity* 24, 189–195.
- Braidwood, R.J., Cambel, H. and Watson, P.J. (1969) Prehistoric investigations in south-western Turkey. *Science* 164, 1275–1276.

- Braidwood, R.J. (1972) Prehistoric investigations in southwestern Asia. *Proc. Am. Phil. Soc.* 116, 310–320.
- Braidwood, L.S., Braidwood, R.J., Howe, B., Reed, C.A. and Watson, P.J. (1983) *Prehistoric archeology along the Zagros flanks*. Oriental Institute Publication 105, University of Chicago Press, Chicago.
- Brandolini, A., Vaccino, P., Boggini, G., Ozkan, H., Kilian, B. and Salamini, F. (2006) Quantification of genetic relationships among A genomes of wheats. *Genome* 49, 297–305.
- Breasted, J.H. (1938) *The conquest of civilization*. Literary Guild of America, New York.
- Breseghele, F. and Sorrells, M.E. (2006) Association mapping of kernel size and milling quality in wheat (*Triticum aestivum* L.) cultivars. *Genetics* 172, 1165–1177.
- Bullrich, L., Appendino, M.L., Tranquilli, G., Lewis, S. and Dubcovsky, J. (2002) Mapping of a thermo-sensitive *earliness per se* gene on *Triticum monococcum* chromosome 1A^m. *Theor. Appl. Genet.* 105, 585–593.
- Bushuk, W. (2001) History, world distribution, production, and marketing. In: W. Bushuk (Ed.), *Rye: Production, Chemistry, and Technology*. American Association of Cereal Chemists, St Paul, Minnesota.
- Cai, W. and Morishima, H. (2002) QTL clusters reflect character associations in wild and cultivated rice. *Theor. Appl. Genet.* 104, 1217–1228.
- Caldwell, K.S., Russell, J., Langridge, P. and Powell, W. (2006) Extreme population-dependent linkage disequilibrium detected in an inbreeding plant species, *Hordeum vulgare*. *Genetics* 172, 557–567.
- Campbell, B., Baenziger, P.S., Gill, K.S., Eskridge, K.M., Budak, H., Erayman, M., Dweikat, I. and Yen, Y. (2003) Identification of QTLs and environmental interactions associated with agronomic traits on chromosome 3A of wheat. *Crop Sci.* 43, 1493–1505.
- Candolle de, A. (1883) (en fait, octobre 1882). *Origine des plantes cultivées*. Germer Baillière, Paris.
- Cao, W., Scoles, G.J. and Hucl, P. (1997) The genetics of rachis fragility and glume tenacity in semi-wild wheat. *Euphytica* 94, 119–124.
- Casas, A.M., Yahiaoui, S., Ciudad, F. and Igartua, E. (2005) Distribution of MWG699 polymorphism in Spanish European barleys. *Genome* 48, 41–45.
- Chandler, P.M., Marion-Poll, A., Ellis, M. and Gubler, F. (2002) Mutants at the *Slender1* locus of barley cv. Himalaya. Molecular and physiological characterization. *Plant Physiol.* 129, 181–190.
- Chantret, N., Cenci, A., Sabot, F., Anderson, O. and Dubcovsky, J. (2004) Sequencing of the *Triticum monococcum* Hardness locus reveals good microcolinearity with rice. *Mol. Genet. Genomics* 271, 377–386.
- Chantret, N., Salse, J., Sabot, F., Rahman, S., Bellec, A., Laubin, B., Dubois, I., Sourdille, P., Joudrier, P., Gautier, M.-F., Cattolico, L., Beckert, M., Aubourg, S., Weissenbach, J., Caboche, M., Bernard, M., Leroy, P. and Chalhou, B. (2005) Molecular basis of evolutionary events that shaped the hardness locus in diploid and polyploid wheat species (*Triticum* and *Aegilops*). *Plant Cell* 17, 1033–1045.
- Chen, Q.F., Yen, C. and Yang, J.L. (1998) Chromosome location of the gene for brittle rachis in the Tibetan weedrace of common wheat. *Genet. Res. Crop. Evol.* 45, 21–25.
- Chikmawati, T., Skovmand, B. and Gustafson, J.P. (2005) Phylogenetic relationships among *Secale* species revealed by amplified fragment length polymorphism. *Genome* 48, 792–801.
- Childe, V.G. (1928) *The most ancient east: The oriental prelude to European prehistory*. Kegan Paul, Trench, Trubner, London.
- Childe, V.G. (1936) *Man makes himself*. Watts, London.
- Cho, C., Kyu, H.O. and Lee, S.H. (1993) Origin, dissemination and utilization of semi-dwarf genes in Korea In: T. Miller, R.M.D. Koebner (Eds.), *Proc. VII Int. Wheat Genetic Symp.* Bath Press, Bath, pp. 223–231.
- Cockram, J., Jones, H., Leigh, F.J., O'Sullivan, D., Powell, W., Laurie, D.A. and Greenland, A.J. (2007) Control of flowering time in temperate cereals: genes, domestication, and sustainable productivity. *J. Exper. Bot.* 58, 1231–1244.

- Damania, A.B. (1998) Diversity of major cultivated plants domesticated in the Near East. In: A.B. Damania, J. Valkoun, G. Willcox, C.O. Qualset, (Eds.), *The origins of agriculture and crop domestication*. Proceedings of the Harlan Symposium. ICARDA, Aleppo, pp. 51–64.
- Danyluk, J., Kane, N.A., Breton, G., Limin, A.E. Fowler, B. and Sarhan, F. (2003) *TaVRT-1*, a putative transcription factor associated with vegetative to reproductive transition in cereals. *Plant Physiol.* 132, 1849–1860.
- Darwin, C. (1859) *On the origin of species by means of natural selection, or the preservation of favored races in the struggle for life*. John Murray, London.
- Darwin, C. (1868) *The variation of animals and plants under domestication*. John Murray, London.
- Davies, M.S. and Hillman, G.C. (1992) Domestication of cereals. In: G.P. Chapman, (Ed.), *Grass evolution and domestication*. Cambridge University Press, Cambridge, pp. 199–244.
- Diamond, J. and Belwood, P. (2003) Farmers and their languages: The first expansions. *Science* 300, 597–603.
- Doebley, J.F., Gaut, B.S. and Smith, B.D. (2006) The molecular genetics of crop domestication. *Cell* 127, 1309–1321.
- Dorofeev, V.F., Filatenko, A.A., Migushova, E.F., Udaczin, R.A. and Jakubziner, M.M. (1979) Wheat. vol. 1. In: V.F. Dorofeev, O.N. Korovina, (Eds.), *Flora of Cultivated Plants*. Leningrad, Russia.
- Dubcovsky, J., Chen, C. and Yan, L. (2005) Molecular characterization of the allelic variation at the *VRN-H2* vernalization locus in barley. *Mol. Breeding* 15, 395–407.
- Dubcovsky, J., Loukoianov, A., Fu, D., Valarik, M., Sanchez, A. and Yan, L. (2006) Effect of photoperiod on the regulation of wheat vernalization genes *VRN1* and *VRN2*. *Plant Mol. Biol.* 60, 469–480.
- Dubcovsky, J. and Dvorak, J. (2007) Genome plasticity a key factor in the success of polyploid wheat under domestication. *Science* 316, 1862–1866.
- Duggan, B., Richards, R.A. and Tsuyuzaki, H. (2002) Environmental effects on the expression of the tiller inhibition (*tin*) gene in wheat. *Funct. Plant. Biol.* 29, 45–53.
- Dunford, R.P., Griffiths, S., Christodoulou, V. and Laurie, D.A. (2005) Characterisation of a barley (*Hordeum vulgare* L.) homologue of the *Arabidopsis* flowering time regulator *GIGANTEA*. *Theor. Appl. Genet.* 110, 925–931.
- Dvorak, J. and Zhang, H.B. (1990) Variation in repeated nucleotide sequences sheds light on the phylogeny of the wheat B and G genomes. *Proc. Natl. Acad. Sci. USA* 87, 9640–9644.
- Dvorak, J. and Luo, M.C. (2001) Evolution of free-threshing and hulled forms of *Triticum aestivum*: old problems and new tools. In: P.D.S. Caligari, P.E. Brandham (Eds.), *The Linnean, Special Issue No 3. Wheat Taxonomy: the legacy of John Percival*. Academic Press, London, pp. 127–136.
- Dvorak, J., Diterlizzi, P., Zhang, H.B. and Resta, P. (1993) The evolution of polyploid wheats: identification of the A genome donor species. *Genome* 36, 21–31.
- Dvorak, J., Luo, M.C. and Yang, Z.L. (1998a) Genetic evidence on the origin of *Triticum aestivum* L. In: A.B. Damania, J. Valkoun, G. Willcox, C.O. Qualset, (Eds.), *The origins of agriculture and crop domestication*. Proceedings of the Harlan Symposium. ICARDA, Aleppo, pp. 235–251.
- Dvorak, J., Luo, M.C., Yang, Z.L. and Zhang, H.B. (1998b) The structure of the *Aegilops tauschii* gene pool and the evolution of hexaploid wheat. *Theor. Appl. Genet.* 67, 657–670.
- Dvorak, J. and Akhunov, E. (2005) Tempos of gene locus deletions and duplications and their relationship to recombination rate during diploid and polyploid evolution in the *Aegilops-Triticum* alliance. *Genetics* 17, 323–332.
- Eastmond, P.J. and Jones, R.L. (2005) Hormonal regulation of gluconeogenesis in cereal aleurone is strongly cultivar-dependent and gibberellin action involves *SLENDER1* but not *GAMYB*. *Plant J.* 44, 483–493.
- Elias, E.M., Steiger, K.D. and Cantrell, R.G. (1996) Evaluation of lines derived from wild emmer chromosome substitutions II. Agronomic traits. *Crop Sci.* 36, 228–233.

- Ellis, M., Rebetzke, G.J., Azanza, F., Richards, R.A. and Spielmeier, W. (2005) Molecular mapping of gibberellin-responsive dwarfing genes in bread wheat. *Theor. Appl. Genet.* 111, 423–430.
- Faris, J. and Gill, B.S. (2002) Genomic targeting and high resolution mapping of the domestication gene *Q* in wheat. *Genome* 45, 706–718.
- Faris, J.D., Fellers, J.P., Brooks, S.A. and Gill, B.S. (2003) A bacterial artificial chromosome contig spanning the major domestication locus *Q* in wheat and identification of a candidate gene. *Genetics* 164, 311–321.
- Faure, S., Higgins, J., Turner, A. and Laurie, D.A. (2007) The *FLOWERING LOCUS T*-like gene family in barley (*Hordeum vulgare*). *Genetics* 176, 599–609.
- Feldman, M. (1966) Identification of unpaired chromosomes in F1 hybrids involving *Triticum aestivum* and *T. timopheevii*. *Can. J. Genet. Cytol.* 8, 144–151.
- Fiedler, H. and Leitner, U. (2000) *Alexander von Humboldts Schriften. Bibliographie der selbständig erschienenen Werke.* (= Beiträge zur Alexander-von-Humboldt-Forschung; 20). Berlin.
- Frederiksen, S. and Petersen, G. (1997) Morphometrical analyses of *Secale* (*Triticeae*, *Poaceae*). *Nordic J. Bot.* 17, 185–197.
- Frederiksen, S. and Petersen, G. (1998) A taxonomic revision of *Secale* (*Triticeae*, *Poaceae*). *Nordic J. Bot.* 18, 399–420.
- Fu, D., Szűcs, P., Yan, L., Helguera, M., Skinner, J.S., Zitzewitz von, J., Hayes, P.M. and Dubcovsky, J. (2005) Large deletions within the first intron in *VRN1* are associated with spring growth habit in barley and wheat. *Mol. Gen. Genomics* 273, 54–65.
- Gandilian, P.A. (1972) On wild growing *Triticum* species of Armenian SSR. *Bot. Zhur.* 57, 173–181.
- Gautier, M., Aleman, M.E., Guirao, A., Marion, D. and Joudrier, P. (1994) *Triticum aestivum* puroindolines, two basic cystine-rich seed proteins: cDNA sequence analysis and developmental gene expression. *Plant Mol. Biol.* 25, 43–57.
- Gautier, M.F., Cosson, P., Guirao, A., Alary, R. and Joudrier, P. (2000) Puroindoline genes are highly conserved in diploid ancestor wheats and related species but absent in tetraploid *Triticum* species. *Plant Sci.* 153, 81–91.
- Gebel, H.G. (2004) There was no centre: the polycentric evolution of the Near Eastern Neolithic. *Neo-lithics* 1/04, 28–32.
- Gepts, P. (1998) What can molecular markers tell us about the process of domestication in common bean? In: A.B. Damania, J. Valkoun, G. Willcox, C.O. Qualset, (Eds.), *The origins of agriculture and crop domestication*. Proceedings of the Harlan Symposium. ICARDA, Aleppo, pp. 198–209.
- Giles, R.J. and Brown, T.A. (2006) *GluDy* allele variations in *Aegilops tauschii* and *Triticum aestivum*: implications for the origins of hexaploid wheats. *Theor. Appl. Genet.* 112, 1563–1572.
- Giroux, M. and Morris, C.G. (1998) Wheat grain hardness results from highly conserved mutations in the friabilin components puroindoline a and b. *Proc. Natl. Acad. Sci. USA* 95, 6262–6266.
- Giroux, M.J., Talbert, L., Habernicht, D.K., Lanning, S., Hempill, A. and Martin, J.M. (2000) Association of puroindoline sequence type and grain hardness in hard red spring wheat. *Crop Sci.* 40, 370–374.
- Goldberg, S.M., Johnson, J., Busam, D., Feldblyum, T., Ferreira, S., Friedman, R., Halpern, A., Khouri, H., Kravitz, S.A., Lauro, F.M., Li, K., Rogers, Y.H., Strausberg, R., Sutton, G., Tallon, L., Thomas, T., Venter, E., Frazier, M. and Venter, J.C. (2006) A Sanger/pyrosequencing hybrid approach for the generation of high-quality draft assemblies of marine microbial genomes. *Proc. Natl. Acad. Sci. USA* 103, 11240–11245.
- Gollan, P., Smith, K. and Bhavé, M. (2007) *Gsp-1* genes comprise a multigene family in wheat that exhibits a unique combination of sequence diversity yet conservation. *J. Cereal Sci.* 45, 184–198.

- Goncharov, N.P., Golovnina, K.A., Kilian, B., Glushkov, S., Blinov, A. and Shumny, V.K. (2008) Evolutionary history of wheats - the main cereal of mankind. In: N. Dobretsov, N. Kolchanov, A. Rozanov, G. Zavarzin (Eds.), *Biosphere Origin and Evolution*. Springer, pp. 407–419.
- Griffiths, S., Dunford, R.P., Coupland, G. and Laurie DA. (2003) The evolution of *CON-STANS*-like gene families in barley, rice and *Arabidopsis*. *Plant Physiol* 131, 1855–1867.
- Gu, Y.Q., Anderson, O.D., Londeorë, C.F., Kong, X., Chibbar, R.N. and Lazo, G.R. (2003) Structural organization of the barley D-hordein locus in comparison with its orthologous regions of wheat genomes. *Genome* 46, 1084–1097.
- Hamblin, M.T., Casa, A.M., Sun, H., Murray, S.C., Paterson, A.H., Aquadro, C.F. and Kresovich, S. (2006) Challenges of detecting directional selection after a bottleneck: Lessons from *Sorghum bicolor*. *Genetics* 173, 953–964.
- Hammer, K. (1984) Das Domestikationssyndrom. *Kulturpflanze* 32, 11–34.
- Hammer, K., Skolimowska, E. and Knüpfper, H. (1987) Vorarbeiten zur monographischen Darstellung von Wildpflanzensortimenten: *Secale L.* *Kulturpflanze* 35, 135–177.
- Hammer, K. (1990) Breeding system and phylogenetic relationships in *Secale L.* *Biol. Zentralbl.* 109, 45–50.
- Harlan, J.R. and Zohary, D. (1966) Distribution of wild wheats and barley. *Science* 153, 1074–1080.
- Harlan, J.R. (1971) Agricultural origins: Centers and noncenters. *Science* 174, 468–474.
- Harlan, J.R. (1975) Our vanishing genetic resources. *Science* 188, 618–621.
- Harlan, J.R. (1995) *The living fields: Our agricultural heritage*. Cambridge University Press, Cambridge.
- Harris, D.R. (1998) The spread of neolithic agriculture from the Levant to western central Asia. In: A.B. Damania, J. Valkoun, G. Willcox, C.O. Qualset, (Eds.), *The origins of agriculture and crop domestication*. Proceedings of the Harlan Symposium. ICARDA, Aleppo, pp. 65–82.
- Haudry, A., Cenci, A., Ravel, C., Bataillon, T., Brunel, D., Poncet, C., Hochu, I., Poirier, S., Santoni, S., Glémin, S. and David, J. (2007) Grinding up wheat: A massive loss of nucleotide diversity since domestication. *Mol. Biol. Evol.* 24, 1506–1517.
- Hawkes, J.G. (1998) Back to Vavilov: Why were plants domesticated in some areas and not in others? In: A.B. Damania, J. Valkoun, G. Willcox, C.O. Qualset, (Eds.), *The origins of agriculture and crop domestication*. Proceedings of the Harlan Symposium. ICARDA, Aleppo, pp. 5–8.
- Hedden, P. (2003) The genes of the Green Revolution. *Trends Genet.* 19, 5–9.
- Heun, M., Schäfer-Pregl, R., Klawan, D., Castagna, R., Accerbi, M., Borghi, B. and Salamini, F. (1997) Site of einkorn wheat domestication identified by DNA fingerprinting. *Science* 278, 1312–1314.
- Hillman, G.C. (1978) On the origins of domestic rye – *Secale cereale*: the finds from Aceramic Can Hasan III in Turkey. *Anatolian Stud.* 28, 157–174.
- Hillman, G.C., Colledge, S.M. and Harris, D.R. (1989) Plant-food economy during the Epipalaeolithic period at Tell Abu Hureyra, Syria: Dietary diversity, seasonality, and modes of exploitation. In: D.R. Harris, G.C. Hillman (Eds.), *Foraging and farming: the evolution of plant exploitation*. Unwin, London, pp. 240–268.
- Hillman, G. and Davies, S. (1990) Measured domestication rates in wild wheats and barley under primitive cultivation, and their archaeological implications. *J. World Prehistory* 4, 157–222.
- Hillman, G. (2000) Plant food economy of Abu Hureyra. In: A. Moore, G. Hillman, T. Legge (Eds.), *Village on the Euphrates, from foraging to farming at Abu Hureyra*. Oxford University Press, New York, pp. 372–392.
- Huang, S., Sirikhachornkit, A., Su, X., Faris, J., Gill, B., Haselkorn, R. and Gornicki, P. (2002) Genes encoding plastid acetyl-CoA carboxylase and 3-phosphoglycerate kinase of the *Triticum/Aegilops* complex and the evolutionary history of polyploidy wheat. *Proc. Natl. Acad. Sci. USA* 99, 8133–8138.

- Humboldt von, A. (1806) Ideen zu einer Geographie der Pflanzen nebst einem Naturgemälde der Tropenländer. Cotta'sche Buchhandlung, Tübingen.
- Iqbal, N., Reader, S.M., Caligari, P.D.S. and Miller, T.E. (2000) The production and characterization of recombination between chromosome 3 N of *Aegilops uniariistata* and chromosome 3A of wheat. *Heredity* 84, 487–492.
- Jaaska, V. (1981) Aspartate aminotransferase and alcohol dehydrogenase isozymes: Intrasppecific differentiation in *Aegilops tauschii* and the origin of the D genome polyploids in the wheat group. *Plant Syst. Evol.* 137, 259–273.
- Jaaska, V. (1995) Isoenzymes in the evaluation of germplasm diversity in wild diploid relatives of cultivated wheat. In: A.B. Damania (Ed.), *Biodiversity and wheat improvement*. John Wiley, Chichester. pp. 247–257.
- Jaaska, V. (1998) On the origin and in statu nascendi domestication of rye and barley: A review. In: A.B. Damania, J. Valkoun, G. Willcox, C.O. Qualset, (Eds.), *The origins of agriculture and crop domestication*. Proceedings of the Harlan Symposium. ICARDA, Aleppo, pp. 210–217.
- Jantasuriyarat, C., Vales, M.I., Watson, C.J.W. and Riera-Lizarazu, O. (2004) Identification and mapping of genetic loci affecting the free-threshing habit and spike compactness in wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* 108, 261–273.
- Johnson, B.L. (1975) Identification of the apparent B-genome donor of wheat. *Can. J. Genet. Cytol.* 17, 21–39.
- Johnson, B.L. and Dhaliwal, H.S. (1976) Reproductive isolation of *Triticum boeoticum* and *Triticum urartu* and the origin of the tetraploid wheats. *Am. J. Bot.* 63, 1088–1094.
- Jones, M.K. (2004) Between Fertile Crescents: Minor grain crops and agricultural origins. In: M.K. Jones (Ed.), *Traces of ancestry: studies in honour of Colin Renfrew*. McDonald Institute for Archaeological Research, Cambridge, pp. 127–135.
- Kanazin, V., Talbert, H., See, D., DeCamp, P., Nevo, E. and Blake, T. (2002) Discovery and assay of single-nucleotide polymorphisms in barley (*Hordeum vulgare*). *Plant Mol. Biol.* 48, 529–537.
- Kan, Y., Wan, Y., Beaudoin, F., Leader, D.J., Edwards, K., Poole, R., Wang, D. Mitchell, R.A.C. and Shewry, P.R. (2006) Transcriptome analysis reveals differentially expressed storage protein transcripts in seeds of *Aegilops* and wheat. *J. Cereal Sci.* 44, 75–85.
- Kandemir, N., Yildirim, A., Kudrna, D.A., Hayes, P.M. and Kleinhofs, A. (2004) Marker assisted genetic analysis of non-brittle rachis trait in barley. *Hereditas* 141, 272–277.
- Kane, N.A., Danyluk, J., Tardif, G., Ouellet, F., Laliberte, J., Limin, A.E., Fowler, D.B. and Sarhan, F. (2005) *TaVRT-2*, a member of the *StMADS-11* clade of flowering repressors, is regulated by vernalization and photoperiod in wheat. *Plant Physiol.* 138, 2354–2363.
- Karsai, I., Szűcs, P., Meszaros, K., Filichkina, T., Hayes, P.M., Skinner, J.S., Lang, L. and Bedő, Z. (2005) The *Vrn-H2* locus is a major determinant of flowering time in a facultative winter growth habit barley (*Hordeum vulgare* L.) mapping population. *Theor. Appl. Genet.* 110, 1458–1466.
- Kato, K., Miura, H., Akiyama, M., Kuroshima, M. and Sawada, S. (1998) RFLP mapping of the three major genes, *Vrn1*, *Q* and *Bl*, on the long arm of chromosome 5A of wheat. *Euphytica* 101, 91–95.
- Kato, K., Miura, H. and Sawada, S. (2000) Mapping QTLs controlling grain yield and its components on chromosome 5A of wheat. *Theor. Appl. Genet.* 101, 933–943.
- Kato, K., Sonokawa, R., Miura, H. and Sawada, S. (2003) Dwarfing effect associated with the threshability gene *Q* on wheat chromosome 5A. *Plant Breed.* 122, 489–492.
- Kerber, E.R. (1964) Wheat: reconstitution of the tetraploid component (AABB) of hexaploids. *Science* 143, 253–255.
- Kerber, E.R. and Dyck, P.L. (1969) Inheritance in hexaploid wheat of leaf rust resistance and other characters derived from *Aegilops squarrosa*. *Can. J. Genet. Cytol.* 11, 639–647.
- Kerber, E.R. and Rowland, G.G. (1974) Origin of the free threshing character in hexaploid wheat. *Can. J. Genet. Cytol.* 16, 145–154.

- Kihara, H. (1924) Cytologische und genetische Studien bei wichtigen Getreidearten mit besonderer Rücksicht auf das Verhalten der Chromosomen und die Sterilität in den Bastarden. Mem. Coll. Sci. Univ. Kyoto Ser. B 1, 1–200.
- Kihara, H. (1944) Discovery of the DD-analyser, one of the ancestors of *Triticum vulgare*. Agric. Hortic. (Tokyo) 19, 13–14.
- Kihara, H., Yamashita, H. and Tanaka, M. (1965) Morphological, physiological, genetical, and cytological studies in *Aegilops* and *Triticum* collected in Pakistan, Afghanistan, Iran. Results of the Kyoto University scientific expedition to the Korakoram and Hindukush in 1955. In: K. Yamashita (Ed.), *Cultivated plants and their relatives*. Kyoto. pp. 4–41.
- Kilian, B., Özkan, H., Kohl, J., Haeseler von, A., Barale, F., Deusch, O., Brandolini, A., Yucel, C., Martin, W. and Salamini, F. (2006) Haplotype structure at seven barley genes: relevance to gene pool bottlenecks, phylogeny of ear type and site of barley domestication. Mol. Gen. Genom. 276, 230–241.
- Kilian, B., Özkan, H., Deusch, O., Effgen, S., Brandolini, A., Kohl, J., Martin, W. and Salamini, F. (2007a) Independent wheat B and G genome origins in outcrossing *Aegilops* progenitor haplotypes. Mol. Biol. Evol. 24, 217–227.
- Kilian, B., Özkan, H., Walther, A., Kohl, J., Dagan, T., Salamini, F. and Martin, W. (2007b) Molecular diversity at 18 loci in 321 wild and 92 domesticate lines reveal no reduction of nucleotide diversity during *Triticum monococcum* (einkorn) domestication: Implications for the origin of agriculture. Mol. Biol. Evol. 24, 2657–2668.
- Kislev, M.E., Nadel, D. and Carmi, I. (1992) Epipalaeolithic (19,000 BP) cereal and fruit diet at Ohalo II, Sea of Galilee. Isr. Rev. Palaeobot. Palynol. 73, 161–166.
- Kislev, M.E. (1980) *Triticum parvicoccum* sp. nov., the oldest naked wheat. Israel J. Bot. 28, 95–107.
- Kislev, M.E. (1984) Botanical evidence for ancient naked wheats in the Near East. In: W. von Zeist and W.A. Casparie (Eds.), *Plants and Ancient Man*. Balkema, Rotterdam, Boston, pp. 141–152.
- Kislev, M. (2002) Origin of annual crops by agro-evolution. Isr. J. Plant Sci. 50, 85–88.
- Kolodinska Brantestam, A., Bothmer von, R., Dayteg, C., Rashal, I., Tuveesson, S. and Weibull, J. (2004) Inter simple sequence repeat analysis of genetic diversity and relationships in cultivated barley of Nordic and Baltic origin. Hereditas 141, 186–192.
- Komatsuda, T., Maxim, P., Senthil, N. and Mano, Y. (2004) High-density AFLP map of nonbrittle rachis 1 (*btr1*) and 2 (*btr2*) genes in barley (*Hordeum vulgare* L.). Theor. Appl. Genet. 109, 986–995.
- Komatsuda, T., Pourkheirandish, M., He, C., Azhaguvel, P., Kanamori, H., Perovic, D., Stein, N., Graner, A., Wicker, T., Tagiri, A., Lundqvist, U., Fujimura, T., Matsuoka, M., Matsumoto, T. and Yano, M. (2007) Six-rowed barley originated from a mutation in a homeodomain-leucine zipper I-class homeobox gene. Proc. Natl. Acad. Sci. USA 104, 1424–1429.
- Kong, X.Y., Gu, Y.Q., You, F.M., Dubcovsky, J. and Anderson, O.D. (2004) Dynamics of the evolution of orthologous and paralogous portions of a complex locus region in two genomes of allploid wheat. Plant Mol. Biol. 54, 55–69.
- Koti, K., Karsai, I., Szücs, P., Horvath, C.S., Meszaros, K., Kiss, G.B., Bedö, Z. and Hayes, P.M. (2006) Validation of the two-gene epistatic model for vernalization response in a winter × spring barley cross. Euphytica 152, 17–24.
- Kranz, A.R. (1963) Die anatomischen, ökologischen und genetischen Grundlagen der Ährenbrüchigkeit des Roggens. Beitr. Biol. Pflanzen. 38, 445–471.
- Kuckuck, H. and Schiemann, E. (1957) Über das Vorkommen von Speltz und Emmer (*Triticum spelta* L. und *T. dicoccum* Schubl.) im Iran. Z. Pflanzenzüchtg. 38, 383–396.
- Kuckuck, H. (1959) Neuere Arbeiten zur Entstehung der hexaploiden Kulturweizen. Z. Pflanzenzüchtg. 41, 205–226.

- Kuraparthi, V., Sood, S. Chunneja, P., Dhaliwal, H.S. and Gill, B.S. (2007) Identification and mapping of tiller inhibition gene (*tin3*) in wheat. *Theor. Appl. Genet.* 114, 285–294.
- Kuraparthi, V., Sood, S. and Gill, B.S. (2008) Genomic targeting and mapping of tiller inhibition gene (*tin3*) of wheat using ESTs and synteny with rice. *Funct. Integr. Genomics* 8, 33–42.
- Ladizinsky, G. (1985) Founder effect in crop-plant evolution. *Econ. Bot.* 39, 191–199.
- Laurie, D., Pratchett, N., Bezant, J.H. and Snape, J.W. (1995) RFLP mapping of five major genes and eight quantitative traits loci controlling flowering time in a winter \times spring barley (*Hordeum vulgare*) cross. *Genome* 38, 575–585.
- Lev-Yadun, S., Gopher, A. and Abbo, S. (2000) The cradle of agriculture. *Science* 288, 1602–1603.
- Li, W. and Gill, B.S. (2006) Multiple genetic pathways for seed shattering in the grasses. *Func. Integr. Genom.* 6, 300–309.
- Li, W., Huang, L. and Gill, B.S. (2008) Recurrent deletions of puroindoline genes at the grain *Hardness* locus in four independent lineages of polyploidy wheat. *Pl. Physiol.* 146, 200–212.
- Lichter, C. (Ed.) (2007) *Die ältesten Monumente der Menschheit*. Badisches Landesmuseum Karlsruhe. Theiss Verlag, Stuttgart.
- Londo, J.P., Chiang, Y.C., Hung, K.H., Chiang, T.Y. and Schaal, B. (2006) Phylogeography of Asian wild rice, *Oryza rufipogon*, reveals multiple independent domestications of cultivated rice, *Oryza sativa*. *Proc. Natl. Acad. Sci. USA* 103, 9578–9583.
- Luo, M.C., Yang, Z.L. and Dvorak, J. (2000) The *Q* locus of Iranian and European spelt wheat. *Theor. Appl. Genet.* 100, 602–606.
- Luo, M., Yang, Z.L., Kota, R.S. and Dvorak, J. (2000) Recombination of chromosomes 3A^m and 5A^m of wheat: the distribution of recombination across chromosomes. *Genetics* 154, 1301–1308.
- Luo, M.C., Yang, Z.L., You, F.M., Kawahara, T., Waines, J.G. and Dvorak, J. (2007) The structure of wild and domesticated emmer wheat populations, gene flow between them, and the site of emmer domestication. *Theor. Appl. Genet.* 114, 947–959.
- Massa, A.N. and Morris (2006) Molecular evolution of the puroindoline-a, puroindoline-b, and grain softness protein-1 genes in the tribe *Triticeae*. *J. Mol. Evol.* 63, 526–536.
- Matsuoka, Y. and Nasuda, S. (2004) Durum wheat as a candidate for the unknown female progenitor of bread wheat: An empirical study with a highly fertile F₁ hybrid with *Aegilops tauschii* Coss. *Theor. Appl. Genet.* 109, 1710–1717.
- MacKey, J. (1954) Neutron and X-ray experiments in wheat and a revision of the speltoid problem. *Hereditas* 40, 65–180.
- McKey, J. (1966) Species relationships in *Triticum*. *Hereditas* 2, 237–276.
- McFadden, E.S. and Sears, E.R. (1946) The origin of *Triticum spelta* and its free-threshing hexaploid relatives. *J. Hered.* 37, 81–89, 107–116.
- Molina-Cano, J.L., Fra-Mon, P., Salcedo, G., Aragoncillo, C., Roca de Togores, F. and Garcia-Olmedo, F. (1987) Morocco as a possible domestication center for barley: Biochemical and agromorphological evidence. *Theor. Appl. Genet.* 73, 531–536.
- Molina-Cano, J.L., Russell, J.R., Moralejo, M.A., Escacena, J.L., Arias, G. and Powell, W. (2005) Chloroplast DNA microsatellite analysis supports a polyphyletic origin for barley. *Theor. Appl. Genet.* 110, 613–619.
- Mori, N., Liu, Y.G. and Tsunewaki, K. (1995) Wheat phylogeny determined by RFLP analysis of nuclear DNA. 2. Wild tetraploid wheats. *Theor. Appl. Genet.* 90, 129–134.
- Mori, N., Ishii, T., Ishido, T., Hirose, S., Watatani, H., Kawahara, T., Nesbitt, M., Belay, G., Takumi, S., Ogihara, Y. and Nakamura, C. (2003) Origin of domesticated emmer and common wheat inferred from chloroplast DNA fingerprinting. 10th International Wheat Genetics Symposium. Paestum, pp. 25–28.
- Morrell, P.L. and Clegg, M.T. (2007) Genetic evidence for a second domestication of barley (*Hordeum vulgare*) east of the Fertile Crescent. *Proc. Natl. Acad. Sci. USA* 104, 3289–3294.
- Morris, C.F. (2002) Puroindolines: the molecular genetic basis of wheat grain hardness. *Plant Mol. Biol.* 48, 633–647.

- Muramatsu, M. (1963) Dosage effect of the *spelta* gene *q* of hexaploid wheat. *Genetics* 48, 469–482.
- Muramatsu, M. (1985) Spike type in two cultivars of *Triticum dicoccum* with the *spelta* gene *q* compared with the *Q*-bearing variety *liguliforme*. *Jpn. J. Breed.* 35, 255–267.
- Muramatsu, M. (1986) The *vulgare* super gene *Q*: its universality in *durum* wheat and its phenotypic effects in tetraploid and hexaploid wheats. *Can. J. Genet. Cytol.* 28, 30–41.
- Nadel, D. (2002) Ohalo II: a 23,000-Year-Old Fisher-Hunter-Gatherer's Camp on the Sea of Galilee. University of Haifa, Haifa.
- Nalam, V., Vales, M.I., Watson, C.J.W., Kianian, S.F. and Riera-Lizarazu, O. (2006) Map-based analysis of genes affecting the brittle rachis character in tetraploid wheat (*Triticum turgidum*). *Theor. Appl. Genet.* 112, 373–381.
- Nalam, V., Vales, M.I., Watson, C.J.W., Johnson, E.B. and Riera-Lizarazu, O. (2007) Map-based analysis of genetic loci on chromosome 2D that affect glume tenacity and threshability, components of the free-threshing habit in common wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* 116, 135–145.
- Nesbitt, M. (1995) Plants and people in ancient Anatolia. *Biblical Archaeologist.* 58, 68–81.
- Nesbitt, M. and Samuel, D. (1996) From stable crop to extinction? The archaeology and history of the hulled wheats. In: S. Padulosi, K. Hammer, J. Heller (Eds.) *Hulled Wheats*. International Plant Genetic Resources Institute, Rome, pp. 41–100.
- Nesbitt, M. (2002) When and where did domesticated cereals first occur in southwest Asia? In: R. Cappers, S. Bottema (Eds.), *The dawn of farming in the Near East*. Berlin. Ex Oriente, pp. 113–132.
- Nishikawa, K., Furuta, Y. and Wada, T. (1980) Genetic studies on alpha-amylase isozymes in wheat. III. Intraspecific variation in *Aegilops squarrosa* and birthplace of hexaploid wheat. *Jpn. J. Genet.* 55, 325–336.
- Onishi, I., Hongo, A., Sasakuma, T., Kawahara, T., Kato, K. and Miura, H. (2006) Variation and segregation for rachis fragility in spelt wheat, *Triticum spelta* L. *Gent. Res. Crop Evol.* 53, 985–992.
- Orabi, J., Backes, G., Wolday, A., Yahyaoui, A. and Jahoor, A. (2007) The horn of Africa as a centre of barley diversification and a potential domestication site. *Theor. Appl. Genet.* 114, 1117–1127.
- Ozkan, H., Brandolini, A., Schaefer-Pregl, R. and Salamini, F. (2002) AFLP analysis of a collection of tetraploid wheats indicates the origin of emmer and hard wheat domestication in southeast Turkey. *Mol. Biol. Evol.* 19, 1797–1801.
- Ozkan, H., Brandolini, A., Pozzi, C., Effgen, S., Wunder, J. and Salamini, F. (2005) A reconsideration of the domestication geography of tetraploid wheats. *Theor. Appl. Genet.* 110, 1052–1060.
- Ozkan, H., Brandolini, A., Torun, A., Altintas, S., Eker, S., Kilian, B., Braun, H., Salamini, F. and Cakmak, I. (2007) Natural variation and identification of microelements content in seeds of einkorn wheat (*Triticum monococcum*). In: H.T. Buck, J.E. Nisi, N. Salomon (Eds.), *Wheat Production in Stressed Environments*. Springer, pp. 455–462.
- Pasternak, R. (1998) Investigations of botanical remains from Nevali Cori PPNB, Turkey: a short interim report. In: A.B. Damania, J. Valkoun, G. Willcox, C.O. Qualset, (Eds.), *The origins of agriculture and crop domestication*. Proceedings of the Harlan Symposium. ICARDA, Aleppo, pp. 170–177.
- Peng, J., Richards, D.E., Hartley, N.M., Murphy, G.P., Devos, K.M., Flintham, J.E., Beales, J., Fish, L.J., Worland, A.J., Pelica, F., Sudhakar, D., Christou, P., Snape, J.W., Gale, M.D. and Harberd, P. (1999) 'Green revolution' genes encode mutant gibberellin response modulators. *Nature* 400, 256–261.
- Peng, J., Ronin, Y., Fahima, T., Röder, M.S., Li, Y., Nevo, E. and Korol, A. (2003) Domestication quantitative trait loci in *Triticum dicoccoides* the progenitor of wheat. *Proc. Natl. Acad. Sci. USA* 10, 2489–2494.

- Perretant, M., Cadalen, T., Charmet, G., Sourdille, P., Nicolas, P., Boeuf, C., Tixier, M.H., Branlard, G., Bernard, S. and Bernard, M. (2000) QTL analysis of bread making quality in wheat using a doubled haploid population. *Theor. Appl. Genet.* 100, 1167–1175.
- Perrino, P., Laghetti, G., D'Antuono, L.F., Al Ajlouni, M., Kanbertay, M., Szabo, A.T. and Hammer, K. (1996) Ecogeographical distribution of hulled wheat species. In: S. Padulosi, K. Hammer, J. Heller (Eds.), *Hulled Wheats*. International Plant Genetic Resources Institute. Rome, pp. 102–118.
- Pluzhnikov, A. and Donnelly, P. (1996) Optimal sequencing strategies for surveying molecular genetic diversity. *Genetics* 144, 1247–1262.
- Pourkheirandish, M. and Komatsuda, T. (2007) The importance of barley genetics and domestication in a global perspective. *Ann. Bot.* 100, 999–1008.
- Pozzi, C., Rossini, L., Vecchietti, A. and Salamini, F. (2004) Gene and genome changes during domestication. In: P.K. Gupta, R.K. Varshney (Eds.), *Cereal genomics*. Kluwer Academic Publisher, London, pp. 165–198.
- Pozzi, C. and Salamini, F. (2007) Genomics of wheat domestication. In: R. Varshney, R. Tuberosa (Eds.), *Genomic assisted crop improvement: Vol 2: Genomics Applications in Crops*. Springer, New York, pp. 441–469.
- Rahman, S., Jolly, J.C., Skerritt, J.H. and Walloscheck, A. (1994) Cloning of a wheat 15 kDa grain softness protein (*GSP*). *GSP* is a mixture of puroindoline-like polypeptides. *Eur. J. Biochem.* 223, 917–925.
- Rao, M. (1972) Mapping of the compactum gene *C* on chromosome 2D of wheat. *Wheat Inf. Serv.* 35, 9.
- Rao, M.V.P. (1977) Mapping of the sphaerococcum gene '*s*' on chromosome 3D of wheat. *Cereal Res. Comm.* 5, 15–17.
- Renfrew, C. (2002) The emerging synthesis: the archaeogenetics of farming/language dispersals and other spread zones. In: P. Bellwood, C. Renfrew (Eds.), *Examining the farming language dispersal hypothesis*. McDonald Institute for Archaeological Research, Cambridge, pp. 3–16.
- Richards, R. (1988) A tiller inhibitor gene in wheat and its effect on plant growth. *Aust. J. Agric. Res.* 39, 749–757.
- Rollefson, G., Simmons, A., Donaldson, M., Gillespie, W., Kafafi, Z., Kohler-Rollefson, I., McAdam, E., Ralston, S. and Tubb, K. (1985) Excavations at the pre-pottery Neolithic B village of 'Ain Ghazal (Jordan), 1983. *Mitteilungen der Deutschen Orient-Gesellschaft zu Berlin* 117, 69–116.
- Rong, J., Millet, E., Manisterski, J. and Feldman, M. (2000) A new powdery mildew resistance gene: introgression from wild emmer into common wheat and RFLP-based mapping. *Euphytica* 115, 121–126.
- Rossini, L., Vecchietti, A., Nicoloso, L., Stein, N., Franzago, S., Salamini, F. and Pozzi, C. (2006) Candidate genes for barley mutants involved in plant architecture: an *in silico* approach *Theor. Appl. Genet.* 112, 1073–1085.
- Russell, J., Booth, A., Fuller, F., Harrower, B., Hedley, P., Machray, G. and Powell, W. (2004) A comparison of sequence-based polymorphism and haplotype content in transcribed and anonymous regions of the barley genome. *Genome* 47, 389–398.
- Saisho, D. and Purugganan, M.D. (2007) Molecular phylogeny of domesticated barley traces expansion of agriculture in the Old World. *Genetics* 177, 1765–1776.
- Sakamura, T. (1918) Kurze Mitteilung über die Chromosomenzahlen und die Verwandtschaftsverhältnisse der *Triticum* Arten. *Bot. Mag. Tokyo* 32, 151–154.
- Salamini, F. (2003) Hormones and the green revolution. *Science* 302, 71–72.
- Salamini, F., Özkan, H., Brandolini, A., Schäfer-Pregl, R. and Martin, W. (2002) Genetics and geography of wild cereal domestication in the Near East. *Nat. Rev. Genet.* 3, 429–441.
- Salamini, F., Heun, M., Brandolini, A., Ozkan, H. and Wunder, J. (2004) Comment on "AFLP data and the origins of domesticated crops." *Genome* 47, 615–620.
- Sax, K. and Sax, M.J. (1924) Chromosome behaviour in a genus cross. *Genetics* 9, 454–464.

- Schiemann, E. (1939) Gedanken zur Genzentrentheorie Vavilovs. *Naturwiss.* 27, 377–401.
- Schiemann, E. and Staudt, G. (1958) *Triticum* \times *dimococcum*, an amphidiploid with genomes AAAABB. *Züchter*. 28, 166–184.
- Schmidt, K. (2001) Göbekli Tepe, southeastern Turkey. A preliminary report on the 1995–1999 excavations. *Paléorient* 26, 45–54.
- Schmidt, K. (2006) *Sie bauten die ersten Tempel*. Verlag CH Beck, München.
- Sears, E.R. (1954) The aneuploids of common wheat. *Res. Bull. Missouri Agric. Exp. Stn.* 572, 1–57.
- Sears, E.R. (1976) A synthetic hexaploid wheat with fragile rachis. *Wheat Info. Serv.* 41/42, 31–32.
- Sencer, H.A. and Hawkes, J.G. (1980) On the origin of cultivated rye. *Biolog J. Linnean Society* 13, 299–313.
- Shao, Q., Li, C. and Basang, C. (1983) Semi-wild wheat from Xizang (Tibet). In: S. Sakamoto (Ed.), *Proceedings of the 6th International Wheat Genetics Symposium, Kyoto, 1983*. Plant Germ-Plasm Institute, Faculty of Agriculture, Kyoto University, Kyoto, Japan, pp. 111–114.
- Shah, M., Gill, K.S., Bezinger, P.S., Yen, Y., Kaeppler, S.M. and Ariyaratne, H.M. (1999) Molecular mapping of loci for agronomic traits on chromosome 3A of bread wheat. *Crop Sci.* 39, 1728–1732.
- Sharma, H. and Waynes, J. (1980) Inheritance of tough rachis in crosses of *Triticum monococcum* and *Triticum boeoticum*. *J. Hered.* 7, 214–216.
- Simonetti, M., Bellomo, M.P., Laghetti, G., Perrino, P., Simeone, R. and Blanco, A. (1999) Quantitative trait loci influencing free-threshing habit in tetraploid wheats. *Gen. Res. Crop Evol.* 46, 267–271.
- Simons, K., Fellers, J.P., Trik, H.N., Zhang, Z., Tai, Y.S., Gill, B.S. and Faris, J.D. (2006) Molecular characterization of the major wheat domestication gene *Q*. *Genetics* 172, 547–555.
- Slageren van, M.W. (1994) Wild wheats: a monograph of *Aegilops* L. and *Amblyopyrum* (Jaub. & Spach) Eig (*Poaceae*). Agriculture University Papers, Wageningen.
- Snape, J., Law, W., Parker, C.N., Worland, B.B. and Worland, A.J. (1985) Genetical analysis of chromosome 5A of wheat and its influence on important agronomic characters. *Theor. Appl. Genet.* 71, 518–526.
- Sogaard, B. and Wettstein-Knowles von, P. (1987) Barley: genes and chromosomes. *Carlsberg Res. Commun.* 52, 123–196.
- Sourdille, P., Perretat, M.R., Charmet, G., Leory, P., Gautire, M.F., Joudrier, P., Nelson, J.C., Sorrells, M.E. and Bernard, M. (1996) Linkage between RFLP markers and genes affecting kernel hardness in wheat. *Theor. Appl. Genet.* 93, 580–586.
- Spielmeier, W. and Richards, R.A. (2004) Comparative mapping of wheat chromosome 1AS which contains the tiller inhibition gene (*tin*) with rice chromosome 5S. *Theor. Appl. Genet.* 109, 1303–1310.
- Stutz, H.C. (1972) On the origin of cultivated rye. *Amer. J. Bot.* 59, 59–70.
- Szücs, P., Karsai, I., Zitzewitz von, J., Meszaros, K., Cooper, L.L., Gu, Y.Q., Chen, T.H., Hayes, P.M. and Skinner, J.S. (2006) Positional relationship between photoperiod response QTL and photoreceptor and vernalization genes in barley. *Theor. Appl. Genet.* 112, 1277–1285.
- Szücs, P., Skinner, J.S., Karsai, I., Cuesta-Marcos, A., Haggard, K.G., Corey, A.E., Chen, T.H.H. and Hayes, P.M. (2007) Validation of the *VRN-H2/VRN-H1* epistatic model in barley reveals that intron length variation in *VRN-H1* may account for a continuum of vernalization sensitivity. *Mol. Genet. Genom.* 277, 249–261.
- Taenzler, B., Esposti, R.F., Vaccino, P., Brandolini, A., Effgen, S., Heun, M., Schäfer-Pregl, R., Borghi, B. and Salamini, F. (2002) Molecular linkage map of Einkorn wheat: mapping of storage-protein and soft-glume genes and bread-making quality QTLs. *Genet. Res. Camb.* 80, 131–143.
- Takahashi, R. (1955) The origin of cultivated barley. In: M. Demerec (Ed.), *Advances in Genetics*. Academic Press, New York, pp. 227–266.

- Takahashi, R. (1972) Non-brittle rachis 1 and non-brittle rachis 2. *Barley Genet. Newsl.* 2, 181–182.
- Takeda, K. (1995) Varietal variation and inheritance of seed dormancy in barley. *Proc. Seventh International Symp on Pre-Harvest Sprouting in Cereals*, Abashiri, Hokkaido, Japan, 205–212.
- Taketa, S., Kikuchi, S., Awayama, T., Yamamoto, S., Ichii, M. and Kawasaki, S. (2004) Monophyletic origin of naked barley inferred from molecular analyses of a marker closely linked to the naked caryopsis gene (*nud*). *Theor. Appl. Genet.* 108, 1236–1242.
- Taketa, S., Awayama, T., Amano, S., Sakurai, Y. and Ichii, M. (2006) High-resolution mapping of the *nud* locus controlling the naked caryopsis in barley. *Plant Breed.* 125, 337–342.
- Taketa, S., Amano, S., Tsujino, Y., Sato, T., Saisho, D., Kakeda, K., Nomura, M., Suzuki, T., Matsumoto, T., Sato, K., Kanamori, H., Kawasaki, S. and Takeda, K. (2008) Barley grain with adhering hulls is controlled by an ERF family transcription factor gene regulating a lipid biosynthesis pathway. *Proc. Natl. Acad. Sci. USA* 105, 4062–4067.
- Talbert, L.E., Smith, L.Y. and Blake, N.K. (1998) More than one origin of hexaploid wheat is indicated by sequence comparison of low-copy DNA. *Genome* 41, 402–407.
- Tanno, K., Taketa, S., Takeda, K. and Komatsuda, T. (2002) A DNA marker closely linked to the *vrs1* locus (row-type gene) indicates multiple origins of six-rowed cultivated barley (*Hordeum vulgare* L.). *Theor. Appl. Genet.* 104, 54–60.
- Tanno, K. and Takeda, K. (2004) On the origin of six-rowed barley with brittle rachis, *agriocrithon* [*Hordeum vulgare* ssp. *vulgare* f. *agriocrithon* (Åberg) Bowd.], based on a DNA marker closely linked to the *vrs1* (six-row gene) locus. *Theor. Appl. Genet.* 110, 145–150.
- Tanno, K. and Willcox, G. (2006) How fast was wild wheat domesticated? *Science* 311, 1886.
- Ternowskaya, T.K. and Zhironov, E.G. (1993) Bread wheat genome D. Genetic control of tender glume and depression at its base. *Tsitologiya Genetica* 27, 78–83.
- Thornsberry, J.M., Goodman, M.M., Doebley, J., Kresovich, S., Nielsen, D. and Buckler, E.S. (2001) *Dwarf8* polymorphisms associate with variation in flowering time. *Nat. Genet.* 28, 286–289.
- Thuillet, A.-C., Bru, D., David, J., Roumet, P., Santoni, S., Sourdille, P. and Bataillon, T. (2002) Direct estimation of mutation rate for 10 microsatellite loci in durum wheat, *Triticum turgidum* (L.) Thell. *Ssp durum* desf. *Mol. Biol. Evol.* 19, 122–125.
- Thuillet, A.-C., Bataillon, T., Poirier, S., Santoni, S. and David, J.L. (2005) Estimation of long-term effective population sizes through the history of durum wheat using microsatellite data. *Genetics* 169, 1589–1599.
- Tranquilli, G., Lijavetzky, D., Muzzi, G. and Dubcovsky, J. (1999) Genetic and physical characterization of grain texture-related loci in diploid wheat. *Mol. Gen. Genet.* 262, 846–850.
- Trevaskis, B., Hemming, M.N., Peacock, W.J. and Dennis, E.S. (2006) *HvVRN2* responds to daylength, whereas *HvVRN1* is regulated by vernalization and developmental status. *Plant Physiol* 140, 1397–1405.
- Turnbull, K., Turner, M., Mukai, Y., Yamamoto, M., Morell, M.K., Appels, R. and Rahman, S. (2003) The organization of genes tightly linked to the *Ha* locus in *Ae. tauschii*, the D genome donor of wheat. *Genome* 46, 330–336.
- Turner, A., Beales, J., Faure, S., Dunford, R.P. and Laurie, D.A. (2005) The pseudo-response regulator *Ppd-H1* provides adaptation to photoperiod in barley. *Science* 310, 1031–1034.
- Vavilov, N.I. (1917) On the origin of cultivated rye. *Bull. Appl. Bot. Genet. Pl. Breed.* 10, 561–590.
- Vavilov, N.I. (1926) *Studies on the origin of cultivated plants*. Institut Botanique Appliqué et d'Amélioration des Plantes, Leningrad.
- Vavilov, N.I. (1992) *Origin and geography of cultivated plants*. (D. Love, transl.), Cambridge University of Press, Cambridge, pp. 316–366.

- Villareal, R., Mujeeb-Kazi, A. and Rajaram, S. (1996) Inheritance of threshability in synthetic hexaploid by *T. aestivum* crosses. *Plant Breed.* 115, 407–409.
- Watanabe, N. and Ikebata, N. (2000) The effects of homoeologous group 3 chromosomes on grain colour dependent seed dormancy and brittle rachis in tetraploid wheat. *Euphytica* 115, 215–220.
- Watanabe, N., Sogiyama, K., Yamagashi, Y. and Skata, Y. (2002) Comparative telosomic mapping of homoeologous genes for brittle rachis in tetraploid and hexaploid wheats. *Hereditas* 137, 180–185.
- Watanabe, N. (2005) The occurrence and inheritance of a brittle rachis phenotype in Italian durum wheat cultivars. *Euphytica* 142, 247–251.
- Watanabe, N., Takesada, N., Fujii, Y. and Martinek, P. (2005a) Comparative mapping of genes for brittle rachis in *Triticum* and *Aegilops* Czech J. Genet. Plant Breed. 41, 39–44.
- Watanabe, N., Takesada, N., Shibata, Y. and Ban, T. (2005b) Genetic mapping of the genes for glaucous leaf and tough rachis in *Aegilops tauschii*, the D-genome progenitor of wheat. *Euphytica* 144, 119–123.
- Watanabe, N., Fujii, Y., Kato, N., Ban, T. and Martinek, P. (2006a) Microsatellite mapping of the genes for brittle rachis on homoeologous group 3 chromosomes in tetraploid and hexaploid wheats. *J. Appl. Genet.* 47, 93–98.
- Watanabe, N., Fujii, Y., Takesada, N. and Martinek, P. (2006b) Cytological and microsatellite mapping of the gene for brittle rachis in a *Triticum aestivum*-*Aegilops tauschii* introgression line. *Euphytica* 151, 63–69.
- Weiss, E., Kislev, M.E. and Hartmann, A. (2006) Autonomous cultivation before domestication. *Science* 312, 1608–1610.
- Wicker, T., Schlagenhauf, E., Graner, A., Close, T.J., Keller, B. and Stein, N. (2006) 454 sequencing put to the test using the complex genome of barley. *BMC Genomics* 7, 275.
- Willcox, G. (1996) Evidence for plant exploitation and vegetation history from three early Neolithic pre-pottery sites on the Euphrates (Syria). *Veget. Hist. Archaeobot.* 5, 143–152.
- Willcox, G. (1998) Archaeobotanical evidence for the beginnings of agriculture in southwest Asia. In: A.B. Damania, J. Valkoun, G. Willcox, C.O. Qualset, (Eds.), *The origins of agriculture and crop domestication*. Proceedings of the Harlan Symposium. ICARDA, Aleppo, pp. 25–38.
- Willcox, G. (2005) The distribution, natural habitats and availability of wild cereals in relation to their domestication in the Near East: Multiple events, multiple centres. *Veget. Hist. Archaeobot.* 14, 534–541.
- Worland, A. (1996) The influence of flowering time genes on environmental adaptability in European wheats. *Euphytica* 89, 49–57.
- Wright, S.I., Vroh, I., Schroeder, S.G., Yamasaki, M., Doebley, J.F., McMullen, M.D. and Gaut, B.S. (2005) The effects of artificial selection on the maize genome. *Science* 308, 1310–1314.
- Yan, L., Loukoianov, A., Tranquilli, G., Helguera, M., Fahima, T. and Dubcovsky, J. (2003) Positional cloning of the wheat vernalization gene *VRN1*. *Proc. Natl. Acad. Sci. USA* 100, 6263–6268.
- Yan, L., Helguera, M., Kato, K., Fukuyama, S., Sherman, J. and Dubcovsky, J. (2004a) Allelic variation at the *VRN-1* promotor region in polyploid wheat. *Theor. Appl. Genet.* 109, 1677–1686.
- Yan, L., Loukoianov, A., Blech, A., Tranquilli, G., Ramakrishna, W., SanMiguel, P., Bennetzen, J.L., Echenique, V. and Dubcovsky, J. (2004b) The wheat *VRN2* gene is a flowering repressor down-regulated by vernalization. *Science* 303, 1640–1644.
- Yan, L., Fu, D., Li, C., Blechl, A., Tranquilli, G., Bonafede, M., Sanchez, A., Valarik, M., Yasuda, S. and Dubcovsky, J. (2006) The wheat and barley vernalization gene *VRN3* is an ortholog of *FT*. *Proc. Natl. Acad. Sci. USA* 103, 19581–19586.
- Zeist van, W. (1970) The oriental institute excavations at Mureybit, Syria: Preliminary report on the 1965 campaign. Part III. Palaeobotany. *J. Near East Stud.* 29, 167–176.

- Zeist van, W. and Roller de, G.J. (1991–2) The plant husbandry of aceramic Cayönü, S.E. Turkey. *Palaeohistorica* 33/34, 65–96.
- Zhang, W., Qu, L.J., Gu, H., Gao, W., Liu, M., Chen, J. and Chen, Z. (2002) Studies on the origin and evolution of tetraploid wheats based on the internal transcribed spacer (ITS) sequences of nuclear ribosomal DNA. *Theor. Appl. Genet.* 104, 1099–1106.
- Zitzewitz von, J., Szűcs, P., Dubcovsky, J., Yan, L., Francia, E., Pecchioni, N., Casas, A., Chen, T.H.H., Hayes, P.M. and Skinner, J.S. (2005) Molecular and structural characterization of barley vernalization genes. *Plant Mol. Biol.* 59, 449–467.
- Zohary, D. (1999) Monophyletic vs. polyphyletic origin of the crops on which agriculture was founded in the Near East. *Genet. Res. Crop Evol.* 46, 133–142.
- Zohary, D. and Hopf, M. (2000) *Domestication of plants in the old world*. Oxford University Press, Oxford.

Chapter 4

Cytogenetic Analysis of Wheat and Rye Genomes

Bikram S. Gill and Bernd Friebe

Abstract Cytogenetics is the correlated study of genetics and cytology. In cereals, five phases of cytogenetic research can be recognized: (i) meiotic pairing analysis of F₁ hybrids; (ii) aneuploidy; (iii) molecular cytogenetics (C-banding and *in situ* hybridization); (iv) deletion bin mapping; and (v) flow cytogenetics. We review here the first four phases of cytogenetic research with special reference to chromosome analysis of wheat and rye. Meiotic pairing analysis revealed genomic relationships among diploid and polyploid species. Aneuploidy opened possibilities of chromosome/arm and comparative mapping. C-banding and *in situ* hybridization allowed rapid identification and analysis of heterochromatic and euchromatic components of wheat and rye chromosomes. The isolation of deletion stocks and their use to study the structure and function of the expressed portion of the wheat genome further revealed structural and functional differentiation of wheat chromosomes into proximal gene-poor/low recombination and distal gene-rich/high recombination compartments. The abovementioned structural and functional differentiation may have been driven by chromosome behavior at meiosis. As DNA sequence information becomes available and with the application of techniques such as Fiber FISH and others that close the gap between DNA level and chromosome level observations, we can truly begin to understand the biological meaning of the superimposed structural, functional, and behavioral differentiation and organization of cereal chromosomes.

4.1 Introduction

Cytogenetics is the correlated study of genetics and cytology including chromosome structure, function, evolution, and behavior during mitosis and meiosis in individuals, populations or hybrids. Cytogenetics also deals with aspects of

B.S. Gill (✉)

Department of Plant Pathology, Kansas State University, Manhattan KS 66506, USA
e-mail: bsgill@ksu.edu

chromosome manipulation for crop improvement and development of special cytogenetic stocks for genome mapping projects. It is convenient to review the cytogenetics of wheat (*Triticum aestivum* L.) and rye (*Secale cereale* L.) in a single chapter as they share six million years of co-evolutionary history (Huang et al. 2002). The first wheat-rye hybrids were made in the early 20th century in Germany (Kattermann 1938). The famous 1B/1R translocation, currently deployed in the world's highest yielding wheat cultivars, was recovered in backcross derivatives based on evolutionary breeding without the benefit of cytogenetics. Wheat has excellent bread-making quality and rye thrives on poor, sandy, and acid soils. The dream of breeders has been to combine these attributes into the man-made crop triticale. Although it has not lived up to its expectations, triticale has found a niche as a forage and animal feed crop adapted to marginal environments. Sears selected the wheat cultivar Chinese Spring as a genetic model because of its ready crossability with rye (Sears and Miller 1985). Wheat-rye addition lines were produced in the 1940s, and wheat chromosomes have been added to the genome of rye (O'Mara 1940; Schlegel 1982). There is a vast amount of cytogenetics literature on these two crops, and this chapter will be a selective review mainly focusing on anchor karyotypes of wheat and rye.

4.2 The Five Phases of Formal Wheat Cytogenetics Research

The formal cytogenetics research era can be divided into five distinct phases and began in 1918 with the determination that $2n=42$ is the correct chromosome number of common wheat (Sakamura 1918). The phase 1 era, from 1918 and lasting into the 1980s, dealt with the meiotic behavior of F_1 hybrids and backcross derivatives. The degree of homology was measured by the ability of homoeologous chromosomes to form chiasmate associations at the metaphase 1 stage of meiosis in F_1 hybrids. Similarly, cytoplasmic homologies were measured by substituting the nucleus of one species into the cytoplasm of another through reciprocal backcrosses. The data on fertility, viability, and vigor of alloplasmic hybrids were used to determine cytoplasmic relationships. Kihara (1919) and Sax (1922) showed in the early 1920s that different wheat species may be diploid, tetraploid, or hexaploid. Based on meiotic pairing, they also showed that diploid and tetraploid species shared one genome (later designated A). Later work showed that the second genome (later designated B) of tetraploid wheat was derived from *Aegilops speltoides* Tausch. They also showed that tetraploid and hexaploid wheat had two genomes (A and B) in common. Kihara (1944) and McFadden and Sears (1946) independently discovered that the third genome of hexaploid wheat was donated by *Ae. tauschii* Coss. Kihara (1954) summarized this enormous work on the genomic relationships of *Triticum* and *Aegilops* species. Kihara's student Tsunewaki in Japan, Maan and Kimber in the U.S., and many others continued research on genomic and

cytoplasmic relationships in the *Triticum/Aegilops* complex. A remarkable symposium was held at Kansas State University in 1994 to celebrate the pending retirements of these three great geneticists where they summarized their lifetime's work (Raupp and Gill 1995).

Phase 2 was the pioneering work of Ernie Sears on wheat aneuploidy begun in the 1930s that ushered in formal genetic analysis in wheat. Sears (1954) isolated monosomics and telosomics for chromosome and arm mapping of genes in wheat. Monosomics were used to produce intervarietal substitution lines for locating and linkage mapping of genes and, later, QTL on individual chromosomes. The aneuploid stocks also were used to determine the genomic affinity of individual chromosomes. Nulli-tetrasomics (Sears 1966) were used to measure chromosome homology based on sporophytic and gemetophytic compensation. O'Mara (1940) produced wheat-rye addition and substitution lines for comparative mapping. Morris and Sears (1967), as well as Sears (1969), reviewed this phase of the cytogenetics research. These approaches overcame the limitations of polyploidy and instead used it as an asset for fine mapping genes and QTL on individual chromosomes and arms in relation to the centromere.

Phase 3 began with the cytogenetic identification of individual wheat and rye chromosomes using the C-banding technique (Gill and Kimber 1974a,b). The biologically meaningful substructures of chromosomes, consisting of heterochromatin and euchromatin, were revealed for the first time and opened to rigorous analysis. Concurrently, biotin-labeled DNA sequences could be directly mapped on chromosomes by *in situ* hybridization on cytological preparations on a glass slide and provided an alternative means of chromosome identification (Rayburn and Gill 1985). Primary trisomics of rye were identified (Zeller et al. 1977). Rye heterochromatin was shown to consist of tandem blocks of satellite DNA repeats (Bedbrook et al. 1980; Appels et al. 1981). Several reviews have appeared on the application of C-banding (Gill and Sears 1988; Friebe and Gill 1996) and *in situ* hybridization techniques (Jiang and Gill 1994, 2006) in plant chromosome and genome analysis.

The isolation of a large number of deletion stocks in common wheat began phase 4 (Endo and Gill 1996). Kihara and collaborators demonstrated that following wheat-*Aegilops* hybridization and backcrossing to wheat, certain *Aegilops* chromosomes were preferentially or exclusively transmitted. Endo (1988) demonstrated that gametes lacking the *Aegilops* chromosome suffered chromosome breakage, including deletions of terminal segments. Molecular markers were used to characterize the deletion stocks and revealed a nonrandom distribution of genes and recombination in wheat (Werner et al. 1992; Gill et al. 1993). Gill and Gill (1994) reviewed this earlier work and suggested a strategy for the mapping and cloning of useful genes in wheat. A large National Science Foundation (NSF) project was initiated to map, based on expressed sequence tag (EST) analysis, a set of wheat unigenes into chromosome bins that were defined by the deletion stocks (<http://wheat.pw.usda.gov/NSF>). The deletion lines used in the NSF-mapping project have been thoroughly characterized (Qi et al. 2003) and widely distributed. The mapping results were published in a

series of papers in a special issue of *Genetics* (168:583–712), and the mapping data summarized by Qi et al. (2004). This large data set was analyzed with respect to the wheat genome, gene organization, and recombination rates (Akhunov et al. 2003a,b) and comparative sequence analysis of the wheat and rice genomes (Sorrells et al. 2003). Recently, Salse et al. (2008) re-analyzed this and performed other comparative sequence analyses with improved statistical parameters to decipher the generation of chromosome diversity in the cereal crops. The EST data on deletion bins anchored to the sequenced rice genome has become a part of the basic tool kit of every wheat geneticist.

The cytogenetic era phase 5, dealing with the application of flow cytometry to wheat aneuploid stocks to isolate individual chromosomes and arms in large quantities to construct BAC libraries, has opened the possibility of sequence ready BAC-contig physical maps (Vrána et al. 2000). This research is reviewed in Chapter 10.

In addition to the reviews relating to aspects of the work discussed above, a large number have dealt with other aspects of wheat cytogenetics literature. Besides general reviews on wheat cytogenetics (Gill 1993; Gill and Friebe 1998), research on wide hybridization, chromosome engineering, and crop improvement has been summarized in a series of papers (Sharma and Gill 1983; Jiang et al. 1994; Friebe et al. 1996; Qi et al. 2007). For the remainder of this chapter, the focus will be on aspects of wheat and rye karyotypic information relevant to the current excitement and interest generated by the wheat genome physical mapping and sequencing projects.

4.3 Wheat Anchor Karyotype

The ideograms of wheat chromosomes to which all genetic and genomic information must be related, and their structure deciphered in terms of DNA sequence composition, have become the common currency of wheat geneticists (Fig. 4.1). It is worth repeating the rules used to establish the standard karyotype and the nomenclature system to describe the 21 individual chromosomes of Chinese Spring wheat (Gill et al. 1991). Prior to this publication, an international chromosome banding nomenclature committee had discussed and approved the rules at the site of the 7th International Wheat Genetics Symposium meetings held in 1988 at Cambridge, UK.

Briefly, band 1 in each arm of a chromosome consists of centromeric heterochromatin followed by a light-staining euchromatic band 2, heterochromatic band 3, and so on, such that all odd-numbered bands consist of heterochromatin and the even-numbered bands are euchromatin (Fig. 4.1). If the total number of bands in each arm exceeds nine, then it is divided into one or more regions, and bands in each region are numbered consecutively. Thus, each chromosome is divisible into biological meaningful heterochromatic and euchromatic regions.

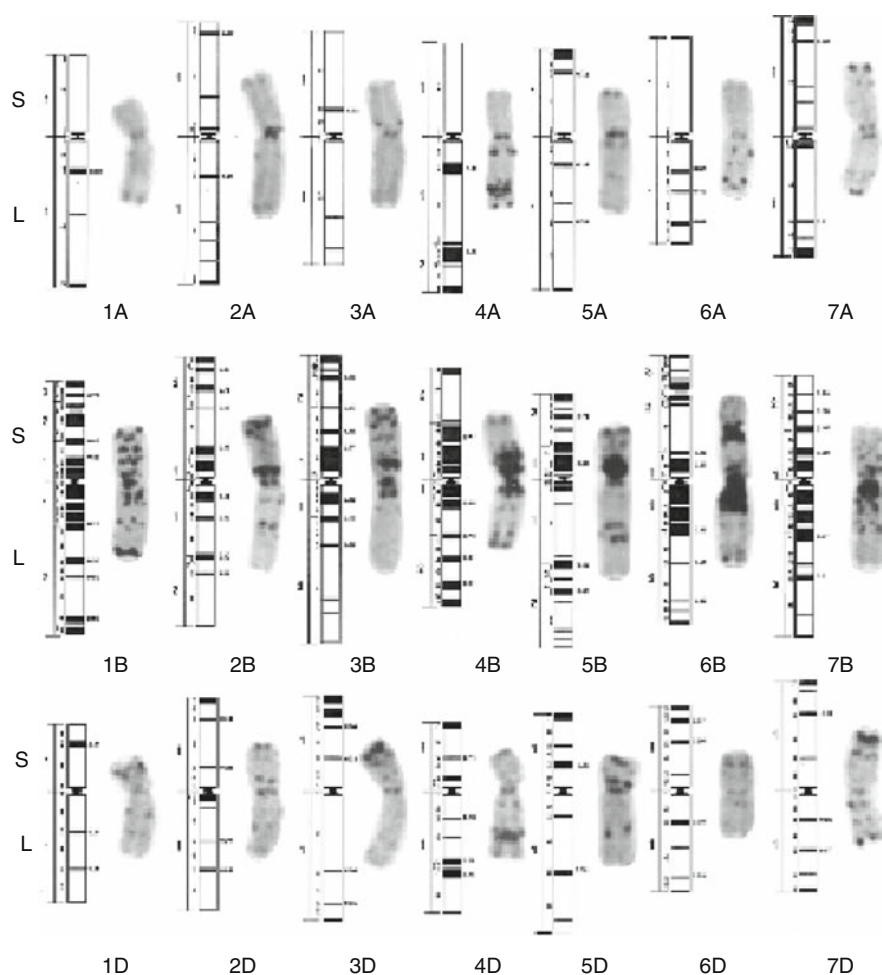


Fig. 4.1 Idiogram representation (*left*) of the C-banded (*right*) patterns of the 21 chromosomes of *Triticum aestivum* L. cv. Chinese Spring. The ideograms are after Gill et al. 1991

As is well established in many organisms, heterochromatin is poor in genes and mainly consists of tandemly repeated DNA families. The GAA-satellite sequence is the major component of the highly heterochromatic B-genome chromosomes. The 120-bp repeated sequence pSc119 that was first isolated in rye (Bedbrook et al. 1980) belongs to the Ty3-gypsy retrotransposon family (Zhang et al. 2004a) and is located in the B-genome heterochromatin (Lapitan et al. 1986; Zhang et al. 2004a,b). pAs1, belonging to the Afa-family (Rayburn and Gill 1986; Nagaki et al. 1998) and another repeat 4P6, a gypsy-like retrotransposon Romani, were localized to the D-genome heterochromatin (Zhang et al. 2004b). A tandem repeat belonging to Nikita LRT-1 element was localized

to the heterochromatic band 4AL13 (see Fig. 4.1) and is conserved in all A-genome diploid and polyploid species (Zhang et al. 2004a). It was localized to the short arm in diploid wheat and the long arm in polyploid wheat thereby providing a cytological marker for the 4A-specific inversion event that occurred during the formation of polyploid wheat.

Other DNA elements belonging to the Afa and CACTA transposon families were located in the subtelomeric regions of all 21 chromosomes of wheat and other Triticeae species (Zhang et al. 2004a; Li et al. 2004). Gypsy-Erika was more predominant in the pericentromeric regions (Li et al. 2004). Seven different, tandemly repeated families, six belonging to the Ty3-gypsy and one to Cereba-like retrotransposons, have been localized to centromeres of cereal chromosomes and vary in sequence, species-specificity, and copy number. The vast majority of the repeats are widely dispersed all over the chromosomes (Zhang et al. 2004a). This brief review underscores the complexity of chromosome organization and the difficulty it will impose on the assembly of physical maps especially for the heterochromatic component of the wheat genome.

4.4 Wheat Chromosome Differentiation

Wheat chromosomes have undergone structural, functional, and behavioral differentiation as revealed by many studies on meiotic behavior, C-banding (heterochromatin polymorphism), and, more recently, by physical mapping data. The first indications of structural differentiation were the observed differences in the amount and pattern of heterochromatin distribution among homoeologous sets of chromosomes that otherwise had similar gene content (Gill and Kimber 1974b; see also Fig. 4.1). This structural polymorphism contributed to differences in chromosome size within homoeologous sets of chromosomes. However, arm ratios were conserved except in the group-4 and group-7 chromosomes (Gill et al. 1991). Chromosome 4B has an arm ratio of 1.1 as compared to 1.7–1.8 in 4A and 4D. In group 7, 7B has an arm ratio of 1.5 as compared to 1.0–1.1 in 7A and 7D.

Other types of structural polymorphisms due to C-banding variation, translocations, and inversions are frequently observed among wheat cultivars (Endo and Gill 1984; Friebe and Gill 1994). Endo and Gill (1984) found that the anomalous arm ratio of chromosome 4B in different wheat cultivars was due to a pericentric inversion; later confirmed by molecular marker data (Mickelson-Young et al. 1995). Most remarkably, meiotic pairing data identified a translocation involving chromosomes 4A and 7B (Gill and Chen 1987; Naranjo et al. 1987). A large portion of the 7BS arm was translocated to 4AL; explaining the anomalous arm ratio of 1.5 in 7B. Molecular data further revealed a pericentric and a paracentric inversion in 4A (Mickelson-Young et al. 1995).

Qi et al. (2006) further studied the dynamic evolution of pericentromeric regions of wheat chromosomes and identified marker signatures for the specific

chromosomal structural alterations. In addition to 4A and 4B, pericentric inversions were detected in chromosomes 2B, 3B, 6B, and 5A. Except for 3B, all arose following polyploidization. A previously documented small translocation involving the tips of 2BS and 6BS in Chinese Spring wheat was confirmed in another winter wheat cultivar.

The isolation of deletion stocks provided additional landmarks for the physical mapping of the wheat genome (Endo and Gill 1996). Over 400 deletion stocks, with a single breakpoint and deletion of the acentric fragment, provide complete coverage of the wheat genome. The massive mapping data revealed functional and behavioral differentiation overlaying the structural organization discussed above. Early deletion-bin mapping papers revealed functional differentiation along the centromere/telomere axis in terms of gradients of gene density and recombination variation (Werner et al. 1992; Gill et al. 1993, 1996a,b; Gill and Gill 1994). This was further confirmed and elaborated upon by the analysis of larger data sets (Akhunov et al. 2003a,b).

Chromosome behavior at meiosis may be partly responsible for the observed functional differentiation. The homologous chromosomes seek out each other during prophase of meiosis for synapsis and chiasmate association in preparation for segregation at anaphase I. Heteromorphic arm combinations, involving a complete arm and an arm with a small terminal deletion, fail to form chiasmate associations (Curtis et al. 1991; Qi et al. 2002, 2005). These observations underscore a key role of the telomeric ends in initiating synapsis at meiosis. Lukaszewski et al. (2004) hypothesized that because of this behavior, terminal ends are synapsed longer and, thus, have a higher probability of recombination compared to more proximal chromosome regions. Recombination rates in turn appear to determine the tempo of new gene evolution (Akhunov et al. 2003a,b; See, Brooks et al. 2006), and this may have produced the observed gradients of gene density and recombination rates along the centromere/telomere axis in wheat. In the Triticeae with large genomes, this discrepancy between physical length and recombination rates will be further distorted compared to cereals with small genomes such as rice.

Most of the agronomically important genes including those for disease resistance are crop-specific and, thus, more recently evolved. Deletion-bin mapping data show that they are invariably located towards the distal ends of chromosomes (Qi et al. 2004; Gill et al. 2008), a fortunate result of the peculiarity of wheat chromosome evolution. Because most of them are located in gene-rich and high-recombination telomeric ends, they are amenable to map-based cloning (Gill and Gill 1994; Qi et al. 2004 and Chapter 12 for additional references).

Chromosome engineering has been an integral part of wheat cytogenetics following Sears' classic experiments on alien gene transfers using x-ray irradiation and induced homoeologous pairing (Sears 1956, 1972, 1973). C-banding and genomic (GISH) and fluorescent (FISH) *in situ* hybridization were powerful tools for determining the location and amount of the alien chromatin introduced into wheat (Friebe et al. 1991; Mukai et al. 1993; also see reviews

by Friebe et al. 1996; Gill et al. 2006). The integration of molecular markers into chromosome-engineering strategies has greatly enhanced the recovery of rare recombinants (Qi et al. 2007). Strategies for enhancing recombination in the proximal regions of wheat chromosomes have been proposed (Qi et al. 2002). Sites of wheat-rye homoeologous recombination appear to be localized (Rogowsky et al. 1993; Lukaszewski et al. 2004). Whole-arm, Robertsonian translocations arise during interkinesis of meiosis II (Friebe et al. 2005) and have hybrid centromeres (Zhang et al. 2001).

4.5 Rye Anchor Karyotype

Even though wheat and rye diverged from a common ancestor only six million years ago (Huang et al. 2002), their karyotypes exhibit dramatic differences in size and structure. The longer rye chromosomes are distinguished easily from wheat chromosomes in wheat \times rye hybrids because the rye genome, at ~ 9 pg, is 33% larger than the size of diploid wheat genome at ~ 6 pg. Rye chromosomes also are distinguished from wheat by the presence of massive blocks of terminal heterochromatin, which was demonstrated by Lima-de-Faria (1952) who made the first high-resolution maps of extended pachytene chromosomes of rye observed at meiotic prophase. Gill and Kimber (1974a) cytogenetically identified the individual rye chromosomes based on their homoeology to wheat chromosomes by C-banding analysis of wheat-rye (cultivar Imperial) addition lines (see also Fig. 4.2). The C-bands observed in somatic chromosomes were equivalent to the heterochromatic structures observed by Lima-de-Faria in pachytene chromosomes. The relative size of the terminal C-bands, presence/absence, and pattern of minor interstitial bands and arm ratios are diagnostic

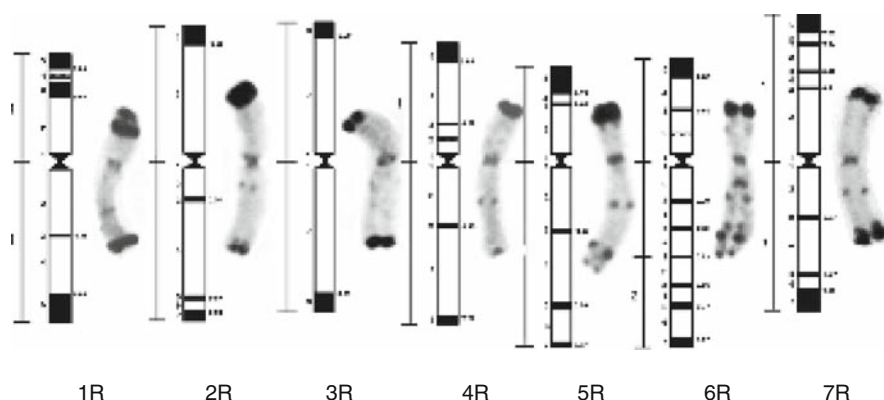


Fig. 4.2 Idiogram representation (left) of the C-banded (right) patterns of the seven chromosomes of *Secale cereale* cv. Imperial. The ideograms are taken from Mukai et al. 1992

for individual chromosome identification. Imperial rye chromosomes are considered to be the standard for arrangement and a nomenclature system was proposed for their description (Sybenga 1983). A comprehensive review of rye cytology, cytogenetics, and genetics was published by Schlegel et al. (1986). Mukai et al. (1992) presented a detailed description of rye karyotypic and molecular cytogenetic information.

C-banding provided a convenient and fast technique and soon demonstrated extensive intraspecific and interspecific polymorphism for heterochromatin in *Secale* species. Bedbrook et al. (1980), Appels et al. (1981), and Appels (1982) took advantage of this polymorphism to clone heterochromatin-specific DNA sequences and provided the first picture of the structure and evolution of heterochromatin in plant chromosomes. Four families of sequences arranged in tandem arrays account for most of the heterochromatin and constitute 12–18% of *S. cereale* DNA. The tandemly organized arrays of 50–60 kb may define the higher-order structure and DNA loop domain regions of rye chromosomes (Vershinin et al. 1995). The repeat sequences are mostly rye-specific and serve as convenient markers for rye chromatin in wheat-rye hybrids (Appels 1982; Lapitan et al. 1986). The 350-bp family in pSc74 was considered specific to rye until its independent amplification was documented in one of the genomes of a perennial grass of the Triticeae (Lapitan et al. 1987). Occasionally, following a deletion event, a repeat family member may be amplified millions of times to the exclusion of other family members (Lapitan et al. 1988).

Comparative mapping revealed that, with the exception of 1R, all other rye chromosomes are rearranged relative to wheat and barley (Devos et al. 1993). Evidently, one or both arms of most of the rye chromosomes have been involved in one or more rounds of translocations. This implies that it will be very difficult to transfer rye genes located in interstitial segments to wheat chromosomes using homoeologous recombination. In such cases, x-ray irradiation may be an alternative method, which was demonstrated for the transfer of the Hessian fly resistance gene *H25* from 6RL to wheat (Friebe et al. 1991; Mukai et al. 1993). Molecular cytogenetic analysis revealed that one of the introgression lines contained less than 1 μ m of rye chromatin at an interstitial location on chromosome 4A of wheat. Restriction fragment length polymorphism mapping data revealed the genetic affinity of the transferred segment to the terminal 6RL segment of rye (Delaney et al. 1995). These papers demonstrated the power of integrated physical and genetic approaches for the fine mapping of alien genes in wheat.

Rye aneuploid stocks, mainly primary trisomics and telotrisomics (Zeller, Kimber, and Gill 1977; see also papers reviewed in Schlegel et al. 1986), have been used for chromosome and arm mapping of genes. Schlegel (1982) even produced monosomic additions of wheat chromosomes to rye. However, because all of these aneuploids do not breed true and are highly sterile, it is difficult to work with them. Instead, single or telosomic additions of rye chromosomes to wheat (Driscoll and Sears 1971; Mukai et al. 1992) have been important material for rye genome mapping. Wheat-rye recombinant

chromosome stocks provide further opportunities for chromosome bin mapping (Rogowsky et al. 1993; Lukaszewski 2000; Lukaszewski et al. 2004). Friebe et al. (2000) and Masoudi-Nejad et al. (2002) have reported the gametocidal factor-induced structural rearrangements in rye chromosomes added to wheat. These include translocations and simple deletions (Fig. 4.3; Friebe et al. 2000) and are an additional, but little exploited, resource for physical mapping of rye chromosomes.

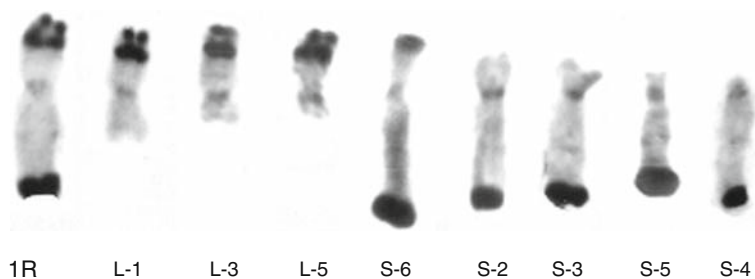


Fig. 4.3 A sample of the deletion chromosomes of rye. The normal chromosome 1R is on the left followed by deletions for the long and short arms. Adapted from Friebe et al. 2000

In this review, we have sampled only a tiny amount of vast literature on rye and very little on triticale cytogenetics. However, work on rye midget chromosomes is worth mentioning with the current interest in plant artificial chromosomes. Alloplasmic wheat in rye cytoplasm leads to lethality, and a genetic factor on 1RL of rye restores vigor and fertility. Japanese workers isolated a 1RL-derived midget chromosome that has been used for molecular cytogenetic mapping (Jackson et al. 1997; Kota et al. 1994) and the isolation of rye-specific centromeric sequences (Francki 2001). Because nuclear-cytoplasmic interactions play such an important role in speciation and agriculture, perhaps this chromosome deserves renewed interest from genome mappers. Friebe et al. (1993) reported a non compensating wheat-rye T3AL·1RL translocation that was maintained in perpetual heterozygosity in alloplasmic wheat.

4.6 Future Prospects

Wheat and rye cytogenetics is the most advanced among plants. We have provided only a small glimpse of the many biologically meaningful cytogenetic observations. However, the lack of DNA sequence information is hampering cereal cytogenetics research. Now that the cost of genome sequencing is falling rapidly and an international wheat genome sequencing consortium (IWGSC, www.wheatgenome.org) have been established, large sequencing data from wheat will soon be available to the cereal community (see Chapter 22). With

sequence in hand, we can truly begin to understand the biological meaning of the superimposed structural, functional, and behavioral differentiation and organization of cereal chromosomes. Techniques such as Fiber FISH, which close the gap between DNA level and chromosome level observations, have been successfully used in wheat for high-resolution mapping (Jackson et al. 2001); these and other more powerful techniques reviewed in Jiang and Gill (2006) will provide an opening for a new era of wheat and Triticeae cytogenetics research.

Acknowledgment Research supported in part by grants from the USDA-CSREES, National Science Foundation, and the Kansas Wheat Commission. This is contribution 08-347-B from the Kansas Agricultural Experiment Station.

References

- Akhunov, E.D., Akhunova, A.R., Linkiewicz, A.M., Dubcovsky, J., Hummel, D., Lazo, G., Chao, S., Anderson, O.D., David, J., Qi, L.L., Echalié, B., Gill, B.S., Miftahudin, Gustafson, J.P., La Rota, M., Sorrells, M.E., Zhang, D., Nguyen, H.T., Kalavacharla, V., Hossain, K., Kianian, S.F., Peng, J., Lapitan, N.L.V., Wennerlind, E.J., Nduati, V., Anderson, J.A., Sidhu, D., Gill, K.S., McGuire, P.E., Qualset, C.O., and Dvorak, J. (2003a). Synteny perturbations between wheat homoeologous chromosomes caused by locus duplications and deletions correlate with recombination rates along chromosome arms. *Proc. Natl. Acad. Sci. USA* 100, 10836–10841.
- Akhunov, E.D., Goodyear, A.W., Geng, S., Qi, L.L., Echalié, B., Gill, B.S., Miftahudin, Gustafson, J.P., Lazo, G., Chao, S., Anderson, O.D., Linkiewicz, A.M., Dubcovsky, J., La Rota, M., Sorrells, M.E., Zhang, D., Nguyen, H.T., Kalavacharla, V., Hossain, K., Kianian, S.F., Peng, J., Lapitan, N.L.V., Gonzalez-Hernandez, J.L., Anderson, J.A., Choi, D-W., Close, T.J., Dilbirli, M., Gill, K.S., Walker-Simmons, M.K., Steber, C., McGuire, P.E., Qualset, C.O., and Dvorak, J. (2003b). The organization and rate of evolution of wheat genomes are correlated with recombination rates along chromosome arms. *Genome Res.* 13, 753–763.
- Appels, R. (1982) The molecular cytology of wheat-rye hybrids. *Int. Rev. Cytol.* 80, 83–132.
- Appels, R., Dennis, E.S., Smyth, D.R., and Peacock, W.J. (1981) Two repeated DNA sequences from the heterochromatic regions of rye chromosomes. *Chromosoma* 70, 265–277.
- Bedbrook, J.R., Jones, J., O'Dell, M., Thompson, R.D., and Flavell, R.B. (1980) A molecular description of telomeric heterochromatin in *Secale* species. *Cell* 19, 545–560.
- Curtis, C.A., Lukaszewski, A.J., and Chrzasiak, M. (1991) Metaphase-I pairing of deficient chromosomes and genetic mapping of deficiency breakpoints in wheat. *Genome* 34, 553–560.
- Delaney, D.E., Friebe, B., Hatchett, J.H., Gill, B.S., and Hulbert, S.H. (1995) Targeted mapping of rye chromatin in wheat by representational difference analysis. *Genome* 38, 458–466.
- Devos, K.M., Atkinson, M.D., Chinoy, C.N., Francis, H.A., Harcourt, R.L., Koebner, R.M. D., Liu, C.J., Masojc, P., Xie, D.X., and Gale, M.D. (1993) Chromosomal rearrangements in the rye genome relative to that of wheat. *Theor. Appl. Genet.* 85, 673–680.
- Driscoll, C.J. and Sears, E.R. (1971) Individual additions of the chromosomes of 'Imperial' rye to wheat. *Agron. Abstr.*, p. 6.
- Endo, T.R. (1988) Induction of chromosomal structural changes by a chromosome of *Aegilops cylindrica* L. in common wheat. *J. Hered.* 79, 366–370.

- Endo, T.R. and Gill, B.S. (1984) Somatic karyotype, heterochromatin distribution, and nature of chromosome differentiation in common wheat, *Triticum aestivum* L. em Thell. *Chromosoma* 89, 361–369.
- Endo, T.R. and Gill, B.S. (1996) The deletion stocks of common wheat. *J. Hered.* 87, 295–307.
- Francki, M.G. (2001) Identification of Bilby, a diverged centromeric Ty1-copia retrotransposon family from cereal rye (*Secale cereale* L.). *Genome* 44, 266–274.
- Friebe, B. and Gill, B.S. (1994) C-band polymorphism and structural rearrangements detected in common wheat (*Triticum aestivum*). *Euphytica* 78, 1–5.
- Friebe, B. and Gill, B.S. (1996) Chromosome banding and genome analysis in diploid and cultivated polyploid wheats. In: P.P. Jauhar (Ed.) *Methods of Genome Analysis in Plants*. CRC Press, Boca Raton, FL, pp. 39–59.
- Friebe, B., Gill, B.S., Mukai, Y., and Maan, S.S. (1993) A noncompensating wheat-rye translocation maintained in perpetual monosomy in alloplasmic wheat. *J. Hered.* 84, 126–129.
- Friebe, B., Hatchett, J.H., Mukai, Y., Gill, B.S., and Sebesta, E.E. (1991) Transfer of Hessian fly resistance from rye to wheat via radiation-induced terminal and intercalary chromosomal translocations. *Theor. Appl. Genet.* 83, 33–40.
- Friebe, B., Jiang, J., Raupp, W.J., McIntosh, R.A., and Gill, B.S. (1996) Characterization of wheat-alien translocations conferring resistance to diseases and pests: current status. *Euphytica* 91, 59–87.
- Friebe, B., Kynast, R.G., and Gill, B.S. (2000) Gametocidal factor-induced structural rearrangements in rye chromosomes added to common wheat. *Chromosome Res.* 8, 501–511.
- Friebe, B., Mukai, Y., Dhaliwal, H.S., Martin, T.J., and Gill, B.S. (1991) Identification of alien chromatin specifying resistance to wheat streak mosaic and greenbug in wheat germ plasm by C-banding and in situ hybridization. *Theor. Appl. Genet.* 81, 381–389.
- Friebe, B., Zhang, P., Linc, G., and Gill, B.S. (2005) Robertsonian translocations in wheat arise by centric misdivision of univalents at anaphase I and rejoining of broken centromeres during interkinesis of meiosis II. *Cytogenet. Genome Res.* 109, 293–297.
- Gill, B.S. (1993) Molecular cytogenetic analysis in wheat. *Crop Sci.* 33, 902–908.
- Gill, B.S. and Chen, P.D. (1987) Role of cytoplasm-specific introgression in the evolution of the polyploid wheats. *Proc. Natl. Acad. Sci. USA* 84, 6800–6804.
- Gill, B.S. and Friebe, B.R. (1998) Plant cytogenetics at the dawn of the 21st century. *Curr. Opin. Pl. Biol.* 1, 109–115.
- Gill, B.S. and Kimber, G. 1974a. The Giemsa C-banded karyotype of rye. *Proc. Natl. Acad. Sci. USA* 10, 1247–1249.
- Gill, B.S. and Kimber, G. 1974b. Giemsa C-banding and the evolution of wheat. *Proc. Natl. Acad. Sci. USA* 10, 4086–4090.
- Gill, B.S. and Sears, R.G. (1988) The current status of chromosome analysis in wheat. In: J.P. Gustafson and R. Appels (Eds.), *Chromosome Structure and Function*. Plenum Press, New York, pp. 299–321.
- Gill, B.S., Friebe, B., and Endo, T.R. (1991) Standard karyotype and nomenclature system for description of chromosome bands and structural aberrations in wheat (*Triticum aestivum*). *Genome* 34, 830–839.
- Gill, B.S., Friebe, B., Raupp, W.J., Wilson, D.L., Cox, T.S., Brown-Guedira, G.L., Sears, R.S., and Fritz, A.K. (2006) Wheat Genetics Resource Center: the first 25 years. *Adv. Agron.* 85, 73–135.
- Gill, B.S., Huang, L., Kuraparthi, V., Raupp, W.J., Wilson, D.L., and Friebe, B. (2008) Alien genetic resources for wheat leaf rust resistance, cytogenetic transfer, and molecular analysis. *Aus. J. Agric. Res.* 59(3), 197–208.
- Gill, K.S. and Gill, B.S. (1994) Mapping in the realm of polyploidy: The wheat model. *BioEssays* 16(11), 841–846.
- Gill, K.S., Gill, B.S., and Endo, T.R. 1993. A chromosome region-specific mapping strategy reveals gene-rich telomeric ends in wheat. *Chromosoma* 102, 374–381.

- Gill, K.S., Gill, B.S., Endo, T.R., and Boyko, E.V. (1996a) Identification and high-density mapping of gene-rich regions in chromosome group 5 of wheat. *Genetics* 143, 1001–1012.
- Gill, K.S., Gill, B.S., Endo, T.R., and Taylor, T. (1996b) Identification and high-density mapping of gene-rich regions in chromosome group 1 of wheat. *Genetics* 144, 1883–1891.
- Huang, S.X., Sirikhachornkit, A., Su, X.J., Faris, J.D., Gill, B.S., Haselkorn, R.G., and Gornicki, P. (2002) Genes encoding plastid acetyl-CoA carboxylase and 3-phosphoglycerate kinase of the *Triticum/Aegilops* complex and the evolutionary history of polyploid wheat. *Proc. Natl. Acad. Sci. USA* 99, 8133–8138.
- Jackson, S.A., Friebe, B., Gill, B.S., and Jiang, J. (1997) Structure of the rye midgut chromosome analyzed by FISH and C-banding. *Genome* 40, 782–784.
- Jackson, S.A., Zhang, P., Chen, W., Phillips, R., Friebe, B., Muthukrishnan, S., and Gill, B.S. (2001) High-resolution structural analysis of biolistic transgene integration into the nuclear genome of wheat. *Theor. Appl. Genet.* 103, 56–62.
- Jiang, J. and Gill, B.S. (1994) Nonisotopic in situ hybridization and plant genome mapping, the first ten years. *Genome* 37, 717–725.
- Jiang, J. and Gill, B.S. (2006) Current status and the future of fluorescence in situ hybridization (FISH) in plant genome research. *Genome* 49, 1057–1068.
- Jiang, J., Friebe, B., and Gill, B.S. (1994) Recent advances in alien gene transfer in wheat. *Euphytica* 73, 199–212.
- Kattermann, G. (1938) Über konstante, halmbehaarte Stämme aus Weizen-Roggen-Bastardierungen mit $2n = 42$ Chromosome. *Z. Ind. Abst. Vererbungs* 1. 74, 354–375.
- Kihara, H. (1919) Über cytologische Studien bei einige Getreidearten. I. Species-Bastarde des Weizens und Weizenroggen-Bastarde. *Bot. Mag. Tokyo* 32, 17–38.
- Kihara, H. (1944) Discovery of the DD-analyser, one of the ancestors of *vulgare* wheats. *Ag. Hort. (Tokyo)* 19, 889–890.
- Kihara, H. (1954) Considerations on the distribution and evolution of *Aegilops* species based on the analyzer method. *Cytologia* 19, 336–357.
- Kota, R.S., Gill, B.S., and Hulbert, S.H. (1994) Presence of various rye-specific repeated DNA sequences on the midgut chromosome of rye. *Genome* 37, 619–624.
- Lapitan, N.L.V., Sears, R.G., Rayburn, A.L., and Gill, B.S. (1986) Wheat-rye translocations. *J. Hered.* 77, 415–419.
- Lapitan, N.L.V., Gill, B.S., and Sears, R.G. (1987) Genomic and phylogenetic relationships among rye and perennial species in the Triticeae. *Crop Sci.* 27, 682–687.
- Lapitan, N.L.V., Sears, R.G., and Gill, B.S. (1988) Amplification of repeated DNA sequences in wheat \times rye hybrids regenerated from tissue culture. *Theor. Appl. Genet.* 75, 381–388.
- Li, W., Zhang, P., Fellers, J.P., Friebe, B., and Gill, B.S. (2004) Sequence composition, organization and evolution of the core Triticeae genome. *The Plant J.* 40, 500–511.
- Lima-de-Faria, A. (1952) Chromomere analysis of the chromosome complement of rye. *Chromosoma* 5, 1–68.
- Lukaszewski, A.J. (2000) Manipulation of the IRS·1BL translocation in wheat by induced homoeologous recombination. *Crop Sci.* 40, 216–225.
- Lukaszewski, A.J., Rybka, K., Korzun, V., Malyshev, S.V., Lapinski, B., and Whitkus, R. (2004) Genetic and physical mapping of homoeologous recombination points involving wheat chromosome 2B and rye chromosome 2R. *Genome* 47, 36–45.
- Masoudi-Nejad, A., Nasuda, S., McIntosh, R.A., and Endo, T.R. (2002) Transfer of rye chromosome segments to wheat by a gametocidal gene. *Chromosome Res.* 10, 349–357.
- McFadden, E.S. and Sears, E.R. (1946) The origin of *Triticum spelta* and its free-threshing hexaploid relatives. *J. Hered.* 37, 81–89.
- Mickelson-Young, L., Endo, T.R., and Gill, B.S. (1995) A cytogenetic ladder-map of wheat homoeologous group-4 chromosomes. *Theor. Appl. Genet.* 90, 1007–1011.
- Morris, R. and Sears, E.R. (1967) The cytogenetics of wheat and its relatives. In: K.S. Quisenberry and L.P. Reitz (Eds.), *Wheat and Wheat Improvement*. Amer. Soc. Agron., Madison, WI, pp. 19–87.

- Mukai, Y., Friebe, B., and Gill, B.S. (1992) Comparison of C-banding patterns and in situ hybridization sites using highly repetitive and total DNA probes of 'Imperial' rye chromosomes added to 'Chinese Spring' wheat. *Jpn. J. Genet.* 67, 71–83.
- Mukai, Y., Friebe, B., Hatchett, J.H., Yamamoto, M., and Gill, B.S. (1993) Molecular cytogenetic analysis of radiation-induced wheat-rye terminal and intercalary chromosomal translocations and the detection of rye chromatin specifying resistance to Hessian fly. *Chromosoma* 102, 88–95.
- Nagaki, K., Tsujimoto, H., and Sasakuma, T. (1998) Dynamics of tandem repetitive Afa-family sequence in Triticeae, wheat-related species. *J. Mol. Evol.* 47, 183–189.
- Naranjo, T., Roca, P.G., Goicoechea, P.G., and Giraldez, R. (1987) Arm homoeology of wheat and rye chromosomes. *Genome* 29, 873–882.
- O'Mara, J.G. (1940) Cytogenetic studies on triticale. I. A method for determining the effects of individual *Secale* chromosomes on *Triticum*. *Genetics* 25, 410–418.
- Qi, L.L., Echalié, B., Friebe, B., and Gill, B.S. (2003) Molecular characterization of a set of wheat deletion stocks for using in chromosome bin mapping of ESTs. *Funct. Integr. Genomics* 3, 39–55.
- Qi, L.L., Echalié, B., Chao, S., Lazo, G.R., Butler, G.E., Anderson, O.D., Akhunov, E.D., Dvorak, J., Linkiewicz, A.M., Ratnasiri, A., Dubcovsky, J., Bermudez-Kandianis, C.E., Greene, R.A., Kantety, R., La Rota, C.M., Munkvold, J.D., Sorrells, S.F., Sorrells, M.E., Dilbirligi, M., Sidhu, D., Erayman, M., Randhawa, H.S., Sandhu, D., Bondareva, S.N., Gill, K.S., Mahmoud, A.A., Ma, X-F., Miftahudin, Gustafson, J.P., Wennerlind, E.J., Nduati, V., Gonzalez-Hernandez, J.L., Anderson, J.A., Peng, J.H., Lapitan, N.L.V., Hossain, K.G., Kalavacharla, V., Kianian, S.F., Pathan, M.S., Zhang, D.S., Nguyen, H.T., Choi, D-W., Close, T.J., McGuire, P.E., Qualset, C.O., and Gill, B.S. (2004) A chromosome bin map of 16,000 expressed sequence tag loci and distribution of genes among the three genomes of polyploid wheat. *Genetics* 168, 701–712.
- Qi, L.L., Friebe, B., and Gill, B.S. (2002) A strategy for enhancing recombination in proximal regions of chromosomes. *Chromosome Res.* 10, 645–654.
- Qi, L.L., Friebe, B., and Gill, B.S. (2005) Origin, structure, and behavior of a highly rearranged deletion chromosome 1BS-4 in wheat. *Genome* 48, 591–597.
- Qi, L.L., Friebe, B., and Gill, B.S. (2006) Complex genome rearrangements reveal evolutionary dynamics of pericentromeric regions in the Triticeae. *Genome* 49, 1628–1639.
- Qi, L.L., Friebe, B., Zhang, P., and Gill, B.S. (2007) Homoeologous recombination, chromosome engineering and crop improvement. *Chromosome Res.* 15, 3–19.
- Raupp, W.J. and Gill, B.S. (Eds.) (1995) *Classical and Molecular Cytogenetic Analysis. Proceedings of a U.S.-Japan Symposium*. Report 95–352-D, Kansas Agricultural Experiment Station, Manhattan, KS.
- Rayburn, A.L. and Gill, B.S. (1985) Use of biotin-labeled probes to map specific DNA sequences on wheat chromosomes. *J. Hered.* 76, 78–81.
- Rayburn, A.L. and Gill, B.S. (1986) Isolation of a D-genome specific repeated DNA sequence from *Aegilops squarrosa*. *Plant Mol. Biol. Rep.* 4, 102–109.
- Rogowsky, P.M., Sorrells, M.E., Shepherd, K.W., and Langridge, P. (1993) Characterisation of wheat-rye recombinants with RFLP and PCR probes. *Theor. Appl. Genet.* 83, 489–494.
- Salse, J., Bolot, S., Throude, M., Jouffe, V., Piegu, B., Quraishi, U.M., Calcagno, T., Cooke, R., Delseny, M., and Feuillet, C. (2008) Identification and characterization of shared duplications between rice and wheat provide new insight into grass genome evolution. *The Plant Cell* 20, 11–24.
- Sakamura, T. (1918) Kurze Mitteilung über die Chromosomenzahlen und die Verwandtschaftsverhältnisse der *Triticum*-Arten. *Bot. Mag. Tokyo* 32, 151–154.
- Sax, K. (1922) Sterility in wheat hybrids. II. Chromosome behavior in partially sterile hybrids. *Genetics* 7, 513–552.
- Schlegel, R. (1982) First evidence for rye-wheat additions. *Biol. Zbl.* 101, 641–646.

- Schlegel, R., Melz, G., and Mettin, D. (1986) Rye cytology, cytogenetics and genetics – Current status. *Theor. Appl. Genet.* 72, 721–734.
- Sears, E.R. (1954) The aneuploids of common wheat. *Mo. Agr. Exp. Sta. Res. Bull.* 572, 1–59.
- Sears, E.R. (1956) The transfer of leaf rust resistance from *Aegilops umbellulata* to wheat. *Brookhaven Symp. Biol.* 9, 1–22.
- Sears, E.R. (1966) Nullisomic-tetrasomic combinations in hexaploid wheat. In: R. Riley and K.R. Lewis (Eds.), *Chromosome manipulations and plant genetics*. Oliver and Boyd, Edinburgh, Scotland, pp. 29–45.
- Sears, E.R. (1969) Wheat cytogenetics. *Ann. Rev. Genet.* 3, 451–468.
- Sears, E.R. (1972) In: *Chromosome engineering in wheat*. Stadler Symposia, University of Missouri, Columbia, 4, 23–38.
- Sears, E.R. (1973) *Agropyron*-wheat transfers induced by homoeologous pairing. In: E.R. Sears and L.M.S. Sears (Eds.), *Proceedings of the 4th International Wheat Genet Symposium*. Agricultural Experiment Station, College of Agriculture, University of Missouri, Columbia, pp. 191–199.
- Sears, E.R. and Miller, T.E. (1985) The history of Chinese Spring wheat. *Cereal Res. Commun.* 13, 261–263.
- See, D.R., Brooks, S.A., Nelson, J.C., Brown-Guedira, G.L., Friebe, B., and Gill, B.S. (2006) Gene evolution at the ends of wheat chromosomes. *Proc. Natl. Acad. Sci. USA* 103, 4162–4167.
- Sharma, H.C. and Gill, B.S. (1983) Current status of wide hybridization in wheat. *Euphytica* 32, 17–31.
- Sorrells, M.E., La Rota, M., Bermudez-Kandianis, C.E., Greene, R.A., Kantety, R., Munkvold, J.D., Miftahudin, Mahmoud, A., Ma, X., Gustafson, J.P., Qi, L.L., Echalié, B., Gill, B.S., Matthews, D.E., Lazo, G.R., Choa, S., Anderson, O.D., Edwards, H., Linkiewicz, A.M., Dubcovsky, J., Akhunov, E.D., Dvorak, J., Zhang, D., Nguyen, H.T., Peng, J., Lapitan, N. L.V., Gonzalez-Hernandez, J.L., Anderson, J.A., Hosssain, K., Kalavacharla, V., Kianian, S.F., Choi, D-W., Close, T.J., Dilbirli, M., Gill, K.S., Steber, C., Walker-Simmons, M.K., McGuire, P.E., and Qualset, C.O. (2003) Comparative DNA sequence analysis of wheat and rice genomes. *Genome Res.* 13, 1818–1827.
- Sybenga, J. (1983) Rye chromosome nomenclature and homoeological relationships. Workshop report. *Z. Pflanzenzucht.* 90, 297–304.
- Vershinin, A.V., Schwarzacher, T., and Heslop-Harrison, P. (1995) The large-scale genomic organization of repetitive DNA families at the telomeres of rye chromosomes. *The Plant Cell* 7, 1823–1833.
- Vrána, J., Kubaláková, M., Simková, H., Cíhalíková, J., Lysák, M.A., and Dolezel, J. (2000) Flow sorting of mitotic chromosomes in common wheat (*Triticum aestivum* L.). *Genetics* 156, 2033–2041.
- Werner, J.E., Endo, T.R., and Gill, B.S. (1992) Toward a cytogenetically based physical map of the wheat genome. *Proc. Natl. Acad. Sci. USA* 89, 11307–11311.
- Zeller, F.J., Kimber, G., and Gill, B.S. (1977) The identification of rye trisomics by translocations and Giemsa staining. *Chromosoma* 62, 279–289.
- Zhang, P., Friebe, B., Lukaszewski, A.J., and Gill, B.S. (2001) The centromere structure in Robertsonian wheat-rye translocation chromosomes indicates that centric breakage-fusion can occur at different positions within the primary constriction. *Chromosoma* 110, 335–344.
- Zhang, P., Li, W., Fellers, J., Friebe, B., and Gill, B.S. (2004a) BAC-FISH in wheat identifies chromosome landmarks consisting of different types of transposable elements. *Chromosoma* 112, 288–299.
- Zhang, P., Li, W., Friebe, B., and Gill, B.S. (2004b) Simultaneous painting of three genomes in hexaploid wheat by BAC-FISH. *Genome* 47, 979–987.

Chapter 5

Applying Cytogenetics and Genomics to Wide Hybridisations in the Genus *Hordeum*

Andreas Houben and Richard Pickering

Abstract Cytogenetic analyses have been widely applied for characterising the barley (*Hordeum vulgare*) genome under the microscope. The methods and results have been extensively reviewed and will only be outlined in this chapter. We focus instead on the applications of cytogenetics and genomics relating to physically mapping the barley genome and for determining species relationships in the genus *Hordeum* through the meiotic analysis of wide hybrids. The application of cytogenetical tools in breeding programmes will be described. Finally we summarise our knowledge of the cellular processes involved in the generation of haploid barley via uniparental chromosome elimination in *H. vulgare* × *H. bulbosum* hybrid embryos.

5.1 Introduction

Barley (*Hordeum vulgare* L.) is one of the major crops cultivated in the temperate zones and the fourth most widely grown small-grain cereal in the world. Besides its agronomic importance, barley has played a significant role in genetic, genomic and cytogenetic studies of the Triticeae. Various aspects of barley cytogenetics including the use of karyotype analysis, chromosomal rearrangements and aneuploidy for mapping and linkage studies have been comprehensively covered previously (Nilan 1964; Ramage 1985; Tsuchiya 1991; Ananiev 1992; Linde-Laursen et al. 1997; Singh 2003; Taketa et al. 2003; Costa and Singh 2006). In this review, therefore, we will summarise attempts to physically map the barley genome and discuss applications of wide barley hybrids using cytogenetic methods in breeding programmes and determining species relationships.

A. Houben (✉)

Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Corrensstrasse 3,
06466 Gatersleben, Germany
e-mail: houben@ipk-gatersleben.de

5.2 Cytological Characterisation and Chromosome Nomenclature of Barley Chromosomes

The genus *Hordeum* is classed in the tribe Triticeae, family Poaceae, which includes several other important crop plant species – wheat (*Triticum aestivum* and *T. turgidum*) and rye (*Secale cereale*) being the most widely grown. There are 32 species and 45 taxa in the genus (Bothmer et al. 1995) and these have been assigned to three genepools based on the concept of Harlan and de Wet (1971). The first genepool comprises *Hordeum vulgare* L. and its closest relative *H. vulgare* subsp. *spontaneum*. The secondary genepool has only one species, *H. bulbosum* L. and the third consists of the remaining species. The genus is quite diverse and there are perennials, annuals, self- and out-pollinating species. The geographical distribution based on centres of diversity is widespread – Southwest Asia, Central Asia, western North America and southern South America (Bothmer et al. 1995). They are generally found in temperate regions of the world, but can occur from the subtropics to the arctic and from sea level to 4500 m a.s.l. The basic chromosome number is 7, and there are diploids ($2n=2\times=14$), tetraploids ($2n=4\times=28$) and hexaploids ($2n=6\times=42$) in the genus. The polyploids have been characterised as allopolyploids, autopolyploids or segmental allopolyploids, and we will look at this in more detail later.

Chromosome sizes across the genus are generally quite similar (Bothmer et al. 1995) and barley itself is diploid with 14 ($2n=2\times=14$, genome formula HH) relatively large chromosomes. The seven chromosomes making up the genome are predominantly metacentric with arm ratios varying from 1 to 1.5 (Fukui and Kakeda 1990; Jensen and Linde-Laursen 1992; Marthe and Künzel 1994). There are two satellited (SAT) chromosomes, i.e. those that carry the nucleolar organising regions (NORs). Based on trisomic studies and genetic marker linkages and relating these to the respective phenotypes, the chromosomes were numbered as 1–5 for non-satellited and 6 and 7 for the satellited chromosomes. The L and S were used to designate the long and short arms, respectively, according to their physical lengths. There was some debate about whether the correct arm was assigned to chromosome 1, since the arm lengths are very similar. In fact, what is now considered the long arm is physically very slightly shorter than its counterpart (Singh and Tsuchiya 1982; Jensen and Linde-Laursen 1992; Marthe and Künzel 1994), but to avoid confusion the original designation has been retained. By utilising wheat-barley chromosome addition lines (Islam et al. 1981) to determine barley chromosome homoeologies with the wheat genome (Hart et al. 1980; Islam et al. 1981; Islam and Shepherd 1992a), the chromosome numbering system was subsequently changed (Linde-Laursen et al. 1997) to become consistent with the Triticeae. Chromosomes 2, 3, 4 and 6 remained the same (2H, 3H, 4H and 6H) whereas 1, 5 and 7 became 7H, 1H and 5H, respectively.

The genome can be identified under the microscope by Giemsa C- and N-banding (e.g. Linde-Laursen 1975). Banding patterns result from staining

of constitutive heterochromatin, which in barley is located mostly near the centromeres and secondary constrictions, with some in the intercalary regions and only a little in the terminal segments. Although minor banding variations are regularly observed, cultivated barley has a common basic chromosome banding pattern that is readily recognisable. Based on a simple and efficient method for the preparation of a high number of barley metaphase spreads suitable for light and electron microscopy (Busch et al. 1996), changes in chromosomal ultrastructure during the cell cycle have been analysed in detail (Martin et al. 1996).

Unlike the genomes of *Arabidopsis thaliana* and rice, the genome of barley is large (1C~5,100 Mbp (Dolezel et al. 1998)) and highly complex with a large number (at least 70–80%) of repetitive sequences (Rimpau et al. 1980). A non-uniform gene distribution in large genomes of cereal crops including barley was suggested earlier by buoyancy density gradient methods (Barakat et al. 1997). Sequence analysis of DNA contigs in barley indicated the presence of clusters of closely linked genes, forming gene islands that are separated by large stretches of repetitive DNA (e.g. Rostoks et al. 2002). To fully exploit the information of the barley genome for crop improvement, sequence analysis of a significantly larger portion of the genome will be needed (Stein 2007). However, the significant content of repetitive DNA and the large genome size will hamper the sequence analysis of the barley genome. One solution to this problem would be to create bacterial artificial chromosome (BAC) libraries specific for small and defined genome segments, similar to those developed for flow-sorted chromosomes of wheat (Safar et al. 2004). Based on available cytogenetic stocks, the barley genome could also be dissected by flow cytometry into sub-fractions (for example chromosome arms), representing only about 6–12% of the total genome (Suchánková et al. 2006) (for more details see Chapter 10).

Various DNA probes (mainly repetitive; listed in Table 5.1) have been used to characterise the barley genome by fluorescence *in situ* hybridisation (FISH). The 45S ribosomal-specific DNA probe pTa71 hybridises to five chromosome pairs (Leitch and Heslop-Harrison 1992); two major sites co-localise with the NORs of chromosomes 5H and 6H. Minor sites are detectable on 1H, 2H and 7H. Sub-telomeric regions of all barley chromosomes can reliably be identified with the barley-specific tandem repeat HvT01 (Belostotsky and Ananiev 1990; Schubert et al. 1998) or the Triticeae-specific AT-rich tandem repeat pHvMWG2315 (Busch et al. 1995). A non-random and motif-dependent distribution of tandem array trinucleotide repeats has been found for barley (Cuadrado and Jouve 2007). With the exception of (ACT)₅, the remaining trinucleotide repeats ((GAA)₅, (AAC)₅, (AAG)₅, (AAT)₅, (AGG)₅, (CAC)₅, (CAT)₅, (CAG)₅, (GCC)₅, (ACG)₅), occur predominantly in the Giemsa-banding positive heterochromatin (Pedersen and Linde-Laursen 1994; Cuadrado and Jouve 2007).

Barley has one of the best-studied Triticeae centromeres, which co-localise with an array of the *cereba* *Ty3/gypsy*-like retroelement (centromeric retroelement in barley) and GC-rich satellite sequences (Presting et al. 1998; Hudakova et al. 2001). Single chromosomes contain an average of about 200 *cereba* elements, amounting to at least 1.4 Mb of centromeric DNA, which are

Table 5.1 DNA probes used for cytological characterisation of barley chromosomes. Minor hybridisation site polymorphisms are possible depending on the barley genotype used

probes	chromosomal locations	comments	References
pTa71 (45S rDNA)	2 major sites, in NORs of 5H and 6H, 3 minor sites on 1H, 2H and 7H	clone derived from wheat (Gerlach and Bedbrook 1979)	(Leitch and Heslop-Harrison 1992)
5S rDNA	2H, 3H, 1H, and 4H, depending on the genotype two additional sites		(Leitch and Heslop-Harrison 1993; Fukui et al. 1994)
<i>Arabidopsis</i> -type telomere repeat	all chromosomal ends		(Schwarzacher and Heslop-Harrison 1991; Wang et al. 1991; Röder et al. 1993)
alpha-amylase-2 gene	long arm 1H	low copy sequence	(Leitch and Heslop-Harrison 1993)
B-hordein locus (Hor2)	distal end of short arm of 1H	low copy sequence	(Lehfer et al. 1993)
GAA-simple sequence repeat	interstitial signals, resembles the C-banding pattern, especially on chromosomes 4H and 2H		(Pedersen et al. 1996; Szakacs and Molnár-Lang 2007)
(CAG) ₅ , (CAC) ₅ , (ACG) ₅ simple sequence repeat	all centromeres except for (CAG) ₅	synthetic oligonucleotides used	(Cuadrado and Jouve, 2007)
(AGG) ₅ , (AAG) ₅ , (AAC) ₅ simple sequence repeat	pericentromeres of all chromosomes	synthetic oligonucleotides used	(Cuadrado and Jouve, 2007)
(ACT) ₅ simple sequence repeat	intercalary, subtelomeric and pericentromeric signals on 2H, 3H, 4H, 5H and 6H	synthetic oligonucleotides used	(Cuadrado and Jouve, 2007)
(CAT) ₅ simple sequence repeat	strong pericentromeric signals on 4H	synthetic oligonucleotides used	(Cuadrado and Jouve, 2007)
BAC7 (Ty3/gypsy-like retroelement cereba and GC-rich satellite)	all centromeres		(Presting et al. 1998; Hudakova et al. 2001)
HvT01 (tandem-repeat)	subtelomeric regions of all chromosomes	No cross-hybridisation with wheat	(Belostotsky and Ananiev, 1990; Schubert et al. 1998; Brandes et al. 1995)

Table 5.1 (continued)

probes	chromosomal locations	comments	References
pHvMWG2315 (AT-rich, tandem repeat)	subtelomeric regions of all chromosomes	also detected in <i>H. vulgare</i> , <i>Agropyron elongatum</i> , <i>S. cereale</i> , <i>T. tauschii</i> , <i>T. turgidum</i> , and <i>T. aestivum</i> .	(Busch et al. 1995)
pHvMWG2314	dispersed signal		(Busch et al. 1995)
pHvA14 (Afa-family repeat)	mainly large subtelomeric regions of all chromosomes and interstitial sites		(Tsujiimoto et al. 1997)
pAs1 (Afa-family repeat)	mainly all subtelomeric regions	clone derived from D genome of wheat (Rayburn and Gill 1986)	(Tsujiimoto et al. 1997)
pSc119.2	sequence is more abundant in wild barley species (Taketa et al. 2000)	clone derived from <i>S. cereale</i> (Bedbrook et al. 1980)	(Xu et al. 1990)

similar to other known grass centromeric retroelements (CR), such as CCS1, CentA, RCB11, RIRE7 or *crwydryn*. Thus, since their species divergence about 60 million years ago (Kumar and Bennetzen 1999), related elements have probably been conserved within the centromere of all cereals because a horizontal transfer between contemporary cereal species is unlikely. Chromatin immunoprecipitation data indicate that CENH3 (a centromere-specific histone H3 that is the assembly site for the kinetochore complex of active centromeres) interacts with barley *cereba* and the GC-rich satellite sequences. As in other organisms, only a portion of the centromeric repeats interacts with CENH3, suggesting that just a small part of the centromeric DNA is utilised in the kinetochore assembly (Houben et al. 2007). The slow mutation rate of centromeric retroelements (Langdon et al. 2000) might be a result of the coevolution of these elements with the interaction partner CENH3. Although this interaction occurs in many eukaryotes, there are stable truncated barley telochromosomes whose centromere does not require the CR *cereba* and GC-satellite sequences to function normally in mitosis and meiosis (Nasuda et al. 2005a). Centromere repeats may, therefore, accumulate after a new centromere has been established.

5.3 Cytogenetics and Species Relationships

Plant morphology, geographical distribution, chromosome morphology, meiotic associations and crossability were traditionally used for assessing species relationships (Morrison and Rajhathy 1959). Karyotypes were based on chromosome morphology, arm ratios and the number of satellite chromosomes (see Section 5.2). Most *Hordeum* species were later characterised according to their Giemsa C-banding patterns (Linde-Laursen et al. 1992 and references therein), which are distributed around the centromere as well as in intercalary and telomeric regions. Some general conclusions were drawn about relationships among *Hordeum* species, but there were anomalies. For example, even though the karyotypes and C-banding patterns of *H. vulgare* and *H. bulbosum* differ (Morrison 1959; Kasha and Sadasivaiah 1971; Linde-Laursen et al. 1990; Taketa et al. 2003), the latter is the most closely related species to cultivated barley outside the primary gene pool.

Refinements to cytogenetics using FISH and chromosome-specific markers have enhanced our ability to identify the chromosome complements of the genus, particularly for *H. vulgare* and *H. bulbosum*. Pedersen et al. (1996) obtained similar patterns to those of C-banding with a short oligonucleotide sequence (GAA)₇. Brown et al. (1999) used several probes including 5S rDNA, 18S rDNA, *Amy-2* cDNA clone and a BAC clone containing the *Hor1* and *Hor2* sequences to identify cultivated barley chromosomes by their discrete signals. This approach was extended by de Bustos et al. (1996) to deduce the relationships and ploidy status (auto- or allopolyploid) of various *Hordeum* species by applying different probes – 5S rDNA (pTa794), 18S-5.8S-26S rDNA

(pTa71) and highly repetitive sequences derived from *Secale cereale* (pSc119.2), *Aegilops tauschii* (pAs1) and *H. chilense* (pHch950). However, it was not possible to assign a chromosome numbering system to most of the wild species' chromosomes.

Using FISH probes ((CTT)₁₀, 5S rDNA, 45S rDNA and pSc119.2) on mitotic and meiotic cells of particular genotypes of *H. bulbosum* and diploid *H. vulgare* × *H. bulbosum* hybrids, *H. vulgare* chromosomes and their respective *H. bulbosum* homoeologues were identified according to their hybridisation signals (Pickering et al. 2006). A major problem is signal polymorphism both within a genotype and among genotypes (Linde-Laursen et al. 1990; de Bustos et al. 1996). For example, the (CTT)₁₀ hybridisation signals in *H. bulbosum* are heteromorphic on chromosomes 2H^bL and 5H^bS (Pickering et al. 2006) and some of the pSc119.2 signals vary in number among accessions as well as being heterozygous within even one plant (de Bustos et al. 1996). This is often found in outpollinating species in *Hordeum* (Bothmer et al. 1995), and makes it difficult to develop a consistent FISH karyotype.

Interspecific genome affinities have also been determined by chromosome associations at meiotic metaphase I (MI) – low frequencies of associations indicating a lack of homoeology between the parental chromosomes, whereas high frequencies might suggest high homoeology. However, there are many inconsistencies and the following points should be considered:

(1) By comparing MI chromosome configurations and chiasma frequency in several polyploid *Hordeum* species with their polyhaploids and interspecific and intergeneric hybrids, the existence of a hemizygous ineffective (haplo-insufficient, Subrahmanyam 1978) or only partially effective (Gupta and Fedak 1985a) control of chromosome pairing was proposed. This was demonstrated by Gupta and Fedak (1985a) in three *Hordeum* species – *H. parodii*, *H. lechleri* and *H. jubatum* – all of which showed predominant bivalent formation at meiosis, with occasional quadrivalents in the latter two species. In polyhaploids of these species about two-thirds of the chromosomes were observed as univalents, with the rest associating mainly as bivalents. Hence, a pairing control mechanism probably exists that is similar to, but more complex than, the *Ph1* gene in hexaploid wheat that normally precludes homoeologous pairing (Riley and Chapman 1958). The mechanism breaks down to some extent in polyhaploids allowing some homoeologues to associate. Confirmation of a pairing control mechanism was obtained from chromosome associations and chiasma frequencies in wide hybrids of these three species with *H. vulgare* and *S. cereale*, in which chromosome associations and chiasmata were suppressed or promoted, respectively (Gupta and Fedak 1985b). Pairing control mechanisms can, therefore, confound data on chromosome affinities.

(2) Differences in chromosome associations can be partly attributed to parental genotype and indicate a genetic control of chromosome pairing. In crosses between *H. vulgare* and *T. aestivum* chiasmata per cell ranged from 0.38 to 3.27 depending on parental genotype (Fedak 1982) and 0.06 in crosses between *H. parodii* (hexaploid) and *H. vulgare* cv 'Betzes' and 7.96 in the

same cross with 'Bomi' (Gupta and Fedak 1985a). In a larger study, Thomas and Pickering (1983) recorded significant differences in chromosome associations and chiasma formation among 74 diploid *H. vulgare* \times *H. bulbosum* hybrids involving ten *H. vulgare* cultivars and three *H. bulbosum* genotypes. Hence, it is hard to draw firm conclusions about species relationships in *Hordeum* when chromosome associations vary so much with parental genotype.

(3) Environmental conditions during meiosis can influence chiasma frequency, chromosome associations and hybrid stability (Bayliss and Riley 1972; Pickering 1990, 1994; Kopyto et al. 1989). However, unless extreme temperatures are used, any minor effects are unlikely to change the overall conclusions from meiotic analyses.

(4) In interspecific and intergeneric crosses involving *Hordeum*, uniparental chromosome elimination often occurs (see detailed section below). Chromosome elimination can, of course, be undesirable when considering species relationships, because both parental sets of chromosomes are a prerequisite for assessing meiotic associations and chiasma frequency.

(5) Most *Hordeum* species have similar sized chromosomes (Bothmer et al. 1995), and although there are differences among species in the morphology of the SAT-chromosomes at mitosis, distinguishing the chromosomes both intraspecifically and in wide hybrids has been very difficult at MI. This is compounded by the chromosomes being very much more condensed at MI than at mitosis. There are a few exceptions, such as rye and barley, whose chromosomes differ greatly in size, allowing parental chromosomes to be identified as univalents or heteromorphic bivalents (Petersen 1991).

(6) Although Giemsa C- or N-banding can differentiate *H. vulgare* chromosomes from each other and chromosomes of other species at mitosis as detailed above, it is problematic at MI because of chromosome condensation and weak signals. Although some barley chromosomes (1H, 4H, 5H and 7H) can be distinguished with some certainty (Xu and Snape 1988), many of the problems associated with identifying chromosomes at mitosis also apply to meiotic preparations and include intraspecific polymorphisms and pattern similarities.

These problems can cause confusion, exemplified by the tetraploid species, *H. depressum*. This species forms regular bivalents at MI (Bothmer et al. 1983, 1987), but a polyhaploid only produced 1.35 bivalents (Pendinen and Chernov 1995). Results from crosses between *H. depressum* \times *H. vulgare* (Morrison and Rajhathy 1959; Bothmer et al. 1983; Pendinen and Chernov 1995), and *H. depressum* \times *S. cereale* (Morrison and Rajhathy 1959; Petersen 1991; Pendinen and Chernov 1995) have also been presented. In crosses with *H. vulgare*, pollen mother cells (PMCs) of the hybrids were quite unstable, with chromosome numbers ranging from 12 to 25, very few bivalents at MI (0.8–3) and some variation among plants. PMCs from hybrids with *S. cereale* were stable (21 chromosomes) and formed more than 6 bivalents (II) at MI (Petersen 1991; Pendinen and Chernov 1995). In contrast, Morrison and Rajhathy (1959) found that chromosome associations were quite similar in *H. depressum* \times *H. vulgare* (1.2II) and *H. depressum* \times *S. cereale* (1.7II) hybrids.

Petersen (1991) attributed the differences in bivalent formation among the various *H. depressum* × *S. cereale* hybrids to dissimilar genetic backgrounds.

Regarding the genome constitution of *H. depressum*, Pendinen and Chernov (1995) proposed that it had evolved from an autopolyploid, but that the genomes had diverged sufficiently to prevent pairing in the polyhaploid and in hybrids with *H. vulgare*. Petersen (1991) suggested that the high bivalent formation in hybrids with *S. cereale* indicated autopolyploidy, but this explanation did not take into account the possibility of a hemizygous ineffective genetic diploidising mechanism that would have allowed homoeologous pairing. All workers agreed that there was no homoeology between *H. depressum* and the other parents, and that *H. vulgare* suppressed chromosome pairing (but cf. Morrison and Rajhathy 1959) whereas *S. cereale* promoted it. Hence, before making conclusions about species relationships, one must consider genotype differences, hybrid stability, suppression or promotion of chromosome associations and genetic diploidising mechanisms.

An interesting method to deduce the genomic composition of a polyploid species is to synthesise it artificially, but this is only possible in a limited number of cases. *H. arizonicum* ($2n=6\times=42$) was re-synthesised successfully by Rajhathy and Symko (1966) after crossing its putative parents (*H. jubatum* ($2n=4\times=28$) and *H. pusillum* ($2n=2\times=14$)). Chromosome numbers of the hybrid were doubled after colchicine treatment and fertility was almost normal. Meiotic chromosome associations in the artificial hybrid were similar to *H. arizonicum*, with predominant bivalent formation, and backcrossing to *H. arizonicum* was successful. There was some multivalent formation in backcross progenies but this declined after selfing and fertility increased. The results indicate that *H. jubatum* and *H. pusillum* (or a close relative of *H. pusillum*, *H. intercedens* ($2n=2\times=14$) – Baum and Bailey 1988) are the most likely parents.

Despite the shortcomings associated with interpretation of meiotic analytical data, wide hybrids have been extremely useful for assessing species relationships in conjunction with other cytogenetic tools and molecular techniques.

5.4 Physical Mapping of the Barley Genome

Genome maps provide primary information for cloning genes or other sequences of interest. Genetic maps have been constructed in barley by using various molecular markers, including for example restriction fragment length polymorphisms (RFLPs) (Graner et al. 1991; Kleinhofs et al. 1993), amplified fragment length polymorphisms (AFLPs) (Qi and Lindhout 1997), simple sequence repeat (SSR)-based markers (Ramsay et al. 2000), expressed sequence tag (EST)-derived SSR markers (Varshney et al. 2006), random amplified microsatellite polymorphisms (RAMPs) (Becker and Heun 1995), and retro-transposon markers (Manninen et al. 2000). However, genetic maps, which are based on recombination values, do not necessarily represent actual physical

location of molecular markers or genes. Such distortions of gene loci can be corrected by physical or cytological mapping. Various cytogenetical methods have been used for the physical mapping of markers or genes. These include PCR on microdissected chromosomes, chromosome deletion mapping and *in situ* hybridisation.

Successful integration of a physical map into genetic maps of barley has become possible by PCR analysis of DNA obtained from various microdissected, cytologically characterised translocation chromosomes (Marthe and Künzel 1994) using primers of genetically mapped DNA sequences (Sorokin et al. 1994; Künzel et al. 2000; Künzel and Waugh 2002). This strategy was highly efficient in relating physical to genetic distances. A very heterogeneous distribution of recombination rates was found along individual chromosomes. Recombination is mainly confined to a few relatively small terminal areas separated by large central segments in which recombination is severely suppressed. The regions of highest recombination frequency (less than or equal to 1 Mb/cM) correspond to only 5% of the total barley genome and harbour around 50% of the 429 markers of the studied RFLP map. The mean physical to genetic ratio in this study was 1.5 Mb/cM in the distal portion of the chromosome arms and 89 Mb/cM near the centromere (Künzel et al. 2000).

Another approach is deletion-based physical mapping using structural changes of a barley chromosome added to wheat (Serizawa et al. 2001). Using addition lines overcame the problems associated with barley's diploid constitution, namely an inability to withstand 'chromosome engineering', which can be tolerated by a polyploid like *T. aestivum*. Chromosomal mutations (deletions and translocations) in barley chromosomes were induced genetically by the gametocidal activity of certain alien chromosomes (e.g. 2C) derived from wild species of the genus *Aegilops*, e.g. *Ae. cylindrica* (Shi and Endo 1999). The induced changes of the barley addition chromosomes were obtained without loss of viability. A set of deletion lines cytologically characterised with GISH, FISH, N- and C-banding enabled the physical allocation of AFLP and STS markers (Serizawa et al. 2001), as well as ESTs to chromosomes 7H and 5H based on PCR amplification using EST-specific primers (Nasuda et al. 2005b; Ashida et al. 2007). The same system was used for the generation of different wheat chromosome deletion lines (Endo 1988) that have been used extensively in the mapping of wheat RFLPs (Werner et al. 1992), AFLPs (Zhang et al. 2000) and ESTs (Randhawa et al. 2004). There is no doubt about the usefulness of physically mapped ESTs and other markers as landmarks in arranging barley BAC clones into contigs. Barley and wheat deletion lines are available from, for example, Japanese National BioResources Project-Wheat (<http://www.shigen.nig.ac.jp/wheat/komugi/top/top.jsp>).

Although *in situ* hybridisation provides the most direct way of physically mapping DNA sequences on chromosomes, *in situ* hybridisation of single or low copy sequences is not yet practical as a routine method for barley chromosomes, but Leitch and Heslop-Harrison (1993) and Lehfer et al. (1993) have determined the physical positions of the low-copy gene families alpha-amylase-2 and B-hordein, respectively. FISH with BAC clones harbouring single and

low-copy sequences provides another approach for saturating genomic regions of interest with markers and building contigs spanning these regions. To suppress genome-wide cross-hybridisation of repeated sequences in BAC clones, excess amounts of unlabeled C₀t-1 fraction of barley DNA have to be added to denatured chromosomes (Lapitan et al. 1997).

More recently, using ultrasensitive FISH (by indirect tyramid signal amplification (TSA)) 14 cDNA probes were assigned to a physical location on one or more chromosomes (Stephens et al. 2004). However, TSA often results in very high background fluorescence because plant cells contain endogenous biotin that also cross-interacts with this signal amplification procedure. It remains to be seen whether directly labelled *in situ* hybridisation probes will improve the visualisation of small DNA targets in barley, as has been demonstrated for maize (Kato et al. 2006). An artificial increase of physical *in situ* hybridisation resolution has been achieved by using super-stretched flow-sorted barley chromosomes as targets. Spatial resolution of neighbouring loci was improved down to 70 kbp as compared with 5-10 Mbp after FISH on mitotic chromosomes (Valarik et al. 2004).

5.5 Generation of Haploid Barley Through Wide Hybridisation and Uniparental Chromosome Elimination

Besides anther culture, interspecific crosses between *H. vulgare* and *H. bulbosum* (bulbous barley grass) are widely used to produce doubled haploid barley plants in breeding programmes (Kasha and Kao 1970; Lange 1971). The hybrids are often karyotypically unstable and the chromosomes of *H. bulbosum* are usually selectively eliminated during early embryonic and endosperm development. The resulting *H. vulgare* haploid embryos are cultured *in vitro* to regenerate haploid plants, which are treated with an aqueous colchicine solution to double the chromosome number and restore fertility. The technique enables large numbers of doubled haploid plants to be obtained for breeding and mapping (Devaux and Pickering 2005). The advantage of doubled haploids for breeders is that homozygosity can be achieved directly from early generation (e.g. F₁) *H. vulgare* hybrid breeding lines, whereas in breeding systems such as pedigree or backcrossing, several selfed generations are needed to attain high levels of homozygosity. Doubled haploids can, therefore, be multiplied for yield trials and agronomic evaluation much earlier than would otherwise be possible with conventional practices for a self-pollinated crop plant. Nevertheless, there are genotype combinations of diploids whose progeny under particular environmental conditions may consist of hybrid plants (Szigat and Pohler 1982; Thomas and Pickering 1983). When it occurs, chromosome elimination is a rapid but irregular process that is limited to early embryonic development. During each somatic cell division, 0–3 chromosomes are eliminated per cell and elimination is completed in embryos 5–9 days after pollination (Subrahmanyam and Kasha 1973; Bennett et al. 1976; Gernand et al. 2006).

Chromosome loss is a step-wise and somewhat variable process that occurs during mitosis regardless of the direction of the cross. The frequency of chromosome elimination is dependent on the genomic constitution of the hybrids: in diploid and tetraploid hybrids in which the parental species contribute equal numbers of genomes, and also in hybrids in which the genome ratio favours the *H. vulgare* parent, elimination is frequent. Conversely, in triploid hybrids (one genome of *H. vulgare* plus two genomes of *H. bulbosum*) elimination is less common (Kasha and Kao 1970; Subrahmanyam and Kasha 1973). The factors causing the elimination of the *H. bulbosum* chromosomes are possibly controlled by genes on chromosomes 2 and 3 of *H. vulgare* (Ho and Kasha 1975) and/or of *H. bulbosum* (Fukuyama and Hosoya 1983). These genes may be balanced by factors on the *H. bulbosum* chromosomes which, when present in sufficient dosage, either neutralise the effect of the *H. vulgare* factors or can 'protect' the *H. bulbosum* chromosomes (Subrahmanyam and Kasha 1973). Chromosome elimination/retention is strongly influenced by parental genotypes and temperature during the early stages of embryo formation; temperatures below 18°C are conducive to chromosome retention (Pickering 1984).

Cytological studies revealed a rapid uniparental chromosome loss by micronuclei formation during mitosis in early hybrid embryos (Bennett et al. 1976; Finch and Bennett 1982; Gernand et al. 2006). Chromosomes destined for elimination were often more peripherally situated on metaphase plates, had smaller centromeric regions, did not segregate properly at metaphase, and lagged behind other chromosomes at anaphase (Fig. 5.1). These observations

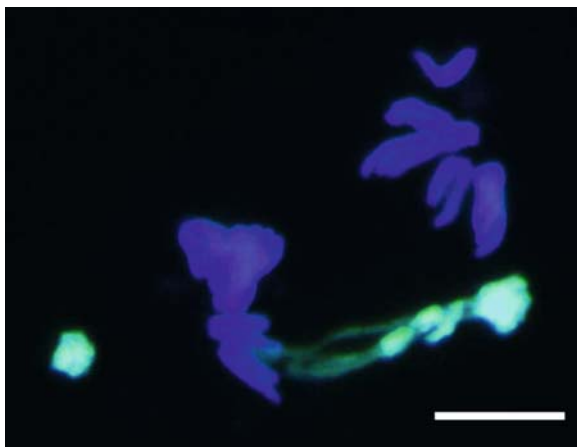


Fig. 5.1 Mitosis-dependent uniparental elimination of *H. bulbosum* chromosomes in *Hordeum vulgare* \times *H. bulbosum* crosses. Anaphase cell of *Hordeum vulgare* \times *H. bulbosum* embryo after *in situ* hybridisation with genomic DNA of *H. bulbosum* (in green). Note asymmetrical segregation of *H. bulbosum* chromosomes. Furthermore, the level of mitotic chromosome condensation partly differed between the parental genomes, with chromosomes of *H. bulbosum* (particularly *H. bulbosum* line Cb3811/5) often less condensed. Bar corresponds to 10 μ m (See Color Insert)

are consistent with the classic mechanism of micronucleus formation, which involves the exclusion of lagging chromosome fragments during re-formation of nuclear membranes at the end of mitosis (Heddle and Carrano 1977; Schubert and Oud 1997; Hudakova et al. 2002).

Several hypotheses have been presented to explain uniparental chromosome elimination during *H. vulgare* × *H. bulbosum* embryo development, e.g. parent-specific suppression of centromere function (Finch and Bennett 1982) possibly due to species-specific organisation of centromeric DNA (Kim et al. 2002), the formation of multipolar spindles, asynchrony of cell division or protein metabolism between the parental genomes (Subrahmanyam and Kasha 1973; Bennett et al. 1976). Lange and Jochensen (1976) speculated that the suppression of the nucleolar activity in the genome of *H. bulbosum* increases its susceptibility to the unknown factor causing chromosome elimination. A novel chromosome elimination pathway was also described (Gernand et al. 2005) that involves the formation of nuclear extrusions during interphase in addition to postmitotically formed micronuclei. It seems that all micronuclei undergo heterochromatinisation as part of the pathway towards haploidisation (Gernand et al. 2006).

5.6 Practical Breeding Applications of Cytogenetics

As a practical tool, cytogenetics has been an invaluable aid for developing new crop plant species. The most important example is triticale (*XTriticosecale*) derived from *Triticum* spp. × *Secale cereale*, but also more recently *XTritordeum* (*Triticum turgidum* conv. *durum* × *H. chilense*; see review by Martín and Cabrera 2005). In the latter case, fertile hybrids were developed that combine the attributes of both parents. Its relevance and importance in agriculture has yet to be realised, but given time and the introduction of more genotypes in breeding programmes to broaden the genetic base, it could provide another way of diversifying crop production. A bonus from the development of *XTritordeum* was a series of addition lines of *H. chilense* chromosomes in *Triticum* backgrounds (Miller et al. 1982; Fernandez and Jouve 1988). *T. turgidum* – *H. chilense* addition lines were initially characterised using isozymes, karyology and morphological traits (Fernandez and Jouve 1988). Analyses with C-banding and FISH signals of *T. aestivum* – *H. chilense* addition lines confirmed the identity of the *H. chilense* chromosomes and enabled a karyotype of *H. chilense* to be prepared (Cabrera et al. 1995).

Another very important series of addition lines was that described by Islam et al. (1981), in which six of the seven possible wheat-barley disomic chromosome addition lines of common wheat were produced. Because of the sterility of telosomic addition for 1HL, only the short arm of 1H is available (Shepherd and Islam 1992; Islam and Shepherd 2000). The addition lines provide basic breeding material for transferring barley genes into wheat.

To obtain the addition lines, barley was used as the female parent, but because of complete sterility the reciprocal cross (wheat \times barley) was favoured despite poor crossability. One 28-chromosome plant was obtained and backcrossed to wheat, and a 49-chromosome progeny plant was subsequently selfed and backcrossed to wheat to get 43- or 44-chromosome offspring. The 43-chromosome lines were partially characterised using isozyme analyses (Hart et al. 1980) and N-banding (Islam 1980) and were classed into six phenotypic groups.

Wheat-barley and wheat-*H. chilense* addition and substitution lines have been extensively used in mapping research, for assigning genetic loci to particular chromosomes (Cho et al. 2006) and to determine the effect of individual chromosomes on phenotype and performance, e.g. brittle rachis (Prieto et al. 2006), heading (Murai et al. 1997) and drought stress (Molnár et al. 2007), as long as the genes on that chromosome are expressed, which is not always the case (Koba et al. 1991).

Another application of addition lines is introgression breeding relying on 'natural' recombination between homoeologues (Malysheva et al. 2003). Wheat-barley hybrids were backcrossed to wheat and putative homoeologous recombination events were observed along with numerical changes in the chromosome constitution. A combination of methods was used for characterising the progenies, including GISH and microsatellite markers. An alternative method for producing recombinant lines is manipulating wheat nullisomic 5B lines (Taketa et al. 2005) or the *ph1b* system (Islam and Shepherd 1992b), which promotes homoeologous pairing. However, the number of recombinants is generally lower than expected (see also Banks et al. 1995; Sherman et al. 2001).

A different approach to introgressing chromatin from wild species is regenerating plants from *in vitro* culture in which non-homoeologous chromosomes recombine following chromosome breakage and fusion in callus tissue prior to plant regeneration. A successful application of this procedure was described by Banks et al. (1995) who identified tissue culture-derived regenerants from wheat-*Thinopyrum* addition lines that were resistant to barley yellow dwarf virus. They selected more resistant plants with this method than by means of adopting the *ph1* approach. More recently, chromosomally engineered wheat-*H. vulgare* plants were regenerated *in vitro* and characterised using combinations of sequential N-banding, GISH and FISH with species-specific probes (Molnár-Lang et al. 2000; Nagy et al. 2002). Molecular markers such as SSR (microsatellites) complemented the cytogenetic tools so that breakpoints of the barley chromosomes making up the translocations could be mapped. Both non-homoeologous and homoeologous translocations were identified and located.

Unfortunately, diploid barley cannot tolerate as much chromosomal manipulation as polyploid wheat. In hybrids between *H. vulgare* and seven *Hordeum* species cultured *in vitro* and with irradiation, plants were identified and characterised cytogenetically with C-banding and conventional Feulgen-staining (Jørgensen and Andersen 1989). However, fertility was very low and no useful

progenies were recovered. With C-banding alone it was impossible to recognise translocations reliably and numerical chromosomal changes were the most frequent aberrations recorded, but some inversions and deletions could also be detected. Regenerants from reciprocal crosses of *H. vulgare* \times *Elymus canadensis* had slightly higher fertility and chromosome pairing than the hybrid from which they were derived, but no data were available on possible translocations between parental chromosomes (Dahleen 1999). Cytoplasmic origin had some influence on chromosome associations.

In interspecific *Hordeum* hybrids, an almost complete set of substitution lines (SLs) of *H. bulbosum* chromosomes into *H. vulgare* was established and characterised by C-banding (Pickering 1992) and RFLP analyses (Timmerman et al. 1993). These were derived from partially fertile triploid 'VBB' hybrids (*H. vulgare* ($2n=2\times=14$) \times *H. bulbosum* ($2n=4\times=28$)) backcrossed to *H. vulgare*. Unfortunately, the SLs were mostly sterile, but from a (7H^b-(7H)) SL, a partially fertile disomic substitution has been obtained (Pickering unpublished). SLs are also a useful mapping tool and for developing species-specific molecular markers.

Introgression lines (ILs) were also produced from backcrosses of 'VBBs' to *H. vulgare* (Pickering et al. 1994) as well as from selfed progeny of partially fertile tetraploid hybrids (VVBB) developed either by colchicine treatment of *H. vulgare* ($2n=2\times=14$) \times *H. bulbosum* ($2n=2\times=14$) diploid hybrids or from *H. vulgare* ($2n=4\times=28$) \times *H. bulbosum* ($2n=4\times=28$) crosses (Szigat and Pohler 1982; Zhang et al. 2001). Progeny plants were screened with a rye repetitive sequence (pSc119.1; McIntyre et al. 1990) in a PCR-based analysis so that only those plants containing *H. bulbosum* DNA were retained (Johnston and Pickering 2002). pSc119.1 hybridises to dispersed sequences across the whole of the *H. bulbosum* genome, but there is little or no cross-hybridisation with the *H. vulgare* genome. Similar proportions of ILs were recovered from backcrossed triploids and selfed tetraploids, but most of the screened tetraploid hybrids were unsuitable because of instability arising from chromosome elimination of the *H. bulbosum* genome (see detailed section below). To date, almost 200 ILs have been fully or partly characterised by GISH and FISH and/or PCR-based chromosome-specific molecular markers for determining into which *H. vulgare* chromosome(s) the *H. bulbosum* chromatin has been integrated (Pickering and Johnston 2005). Introgressed segments are usually distally located on the *H. vulgare* chromosomes and up to three have been found in individual ILs. Their physical sizes vary greatly (Fig. 5.2) and can be related to genetic size by comparing GISH data with molecular mapping. Their agronomic performance depends on introgression size and location; this was tested in a split yield trial in which disease was controlled with fungicides or allowed to develop. Seven of the 13 ILs tested had yield increases ranging from 4 to 22% more than the recurrent barley parents when disease was present, whereas a yield penalty of up to 36% was associated with seven ILs when disease was controlled (Pickering et al. 2004). Taking a specific example, an IL with a very small introgression on chromosome 2HL (Pickering et al. 2000) conferring resistance to leaf rust (*Puccinia hordei* Otth) had a similar yield to its barley parent in the absence of leaf rust. When leaf rust was present, the IL yielded 14% higher.

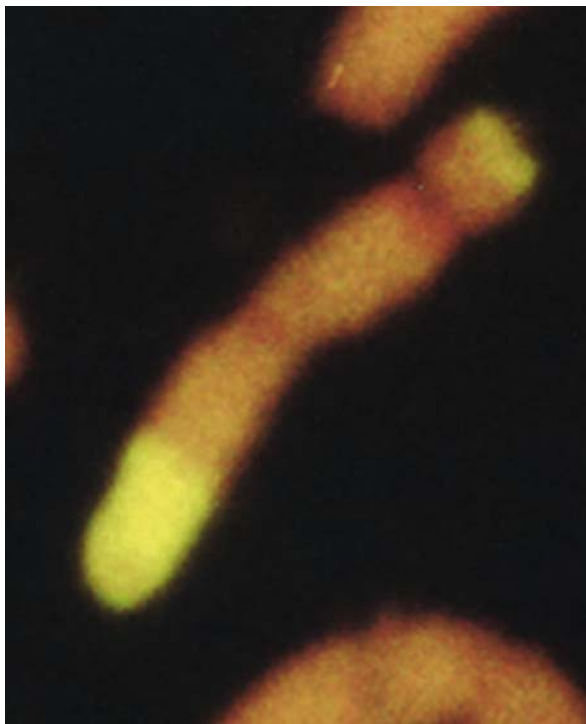


Fig. 5.2 Genomic *in situ* hybridisation of chromosome 6H derived from a *Hordeum vulgare* – *H. bulbosum* introgression line. Two introgressions (yellow signals) of *Hordeum bulbosum* chromatin can be seen on *H. vulgare* chromosome 6H. Note the larger size of the introgression on 6HL compared with the small distal introgression on 6HS (See Color Insert)

Current numbers of ILs with introgressions located on different chromosomes range from 2 (chromosome 3HL) to 48 (2HL) (Fig. 5.3). More introgressions tend to be found on the long arms of individual chromosomes (Fig. 5.3), which is to be expected because the greater physical size allows more chance of homoeologous pairing and recombination. That chromosome associations occur more frequently among the long than the short arms was confirmed when data were collated from a diploid ‘VB’ *H. vulgare* \times *H. bulbosum* hybrid using FISH markers (Fig. 5.4). However, other factors such as recombination hotspots (Künzel et al. 2000) must also be considered. It is interesting that the lowest frequency of chromosome associations occurred between 5HS–5H^bS. 5HS, which contains one of the NORs, has the fewest intermediate and high recombination hotspots in barley (Künzel et al. 2000). Reduced short-arm associations between 5H barley homologues using FISH (Pickering et al. 2005) and ditelosomic 5HS additions in wheat compared with other chromosome arms have been reported (Islam 1983). King et al. (2007) also recorded reduced recombination associated with the NOR in a *Lolium perenne* – *Festuca pratensis* chromosome 3 substitution.

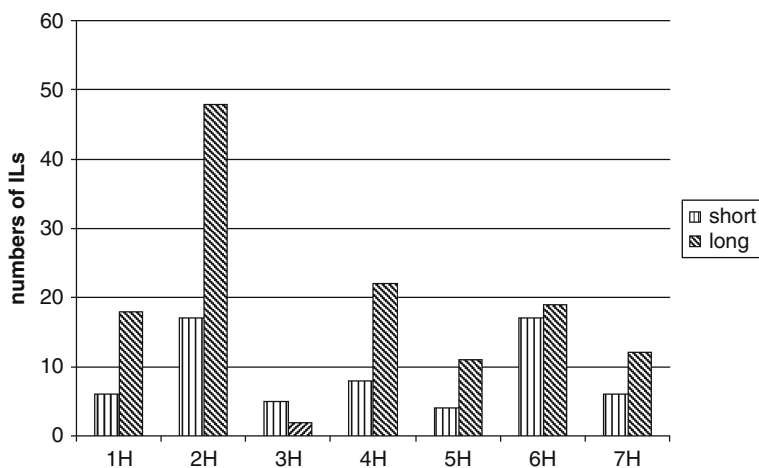


Fig. 5.3 Numbers of characterised introgression lines (ILs) that contain DNA from *Hordeum bulbosum* introgressed into individual barley chromosome arms

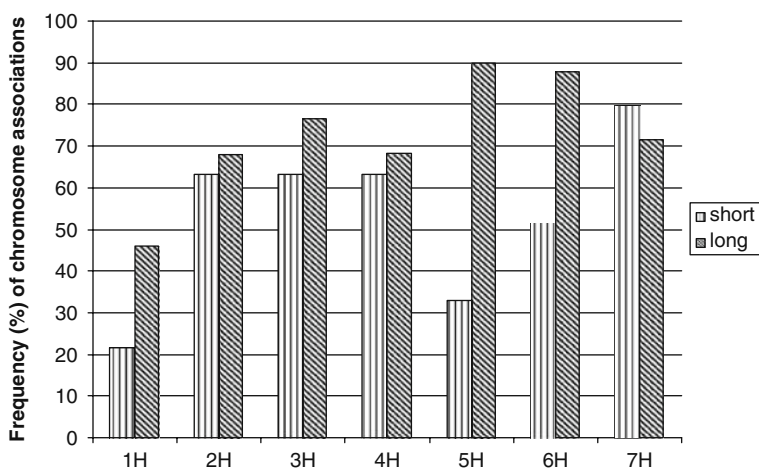


Fig. 5.4 Frequency (%) of homoeologous *Hordeum vulgare* - *H. bulbosum* chromosome arm associations in PMCs derived from two diploid *Hordeum vulgare* × *H. bulbosum* hybrids. (Modified from Table 1 in Pickering et al. 2006 and reprinted with permission of 'Genome')

In addition to the leaf rust resistance on 2HL mentioned above, resistances to several other diseases have been identified that are associated with introgressions on specific chromosomes (Xu and Kasha 1992; Walther et al. 2000; Ruge et al. 2003; Pickering and Johnston 2005). More recently, a novel source of tolerance to barley yellow dwarf virus was transferred to *H. vulgare* chromosome 2HL from *H. bulbosum* (Pickering and Habekuss unpublished). *Vrn* loci

and increased tolerance to low temperatures have also been introgressed from *H. bulbosum* and mapped to interstitial introgressions on chromosome 5HL (Balko, Rizza and Pickering unpublished). Another introgression line (38U16) with a large introgression on 5HL (Fig. 5.5) appears to possess properties similar to the gametocidal gene in wheat-*Ae. cylindrica* addition lines (see details earlier in the text). In the case of 38U16, almost 100% of selfed progenies from an F₁ introgression heterozygote (38U16 × Golden Promise) were homozygous for the introgression. Seed setting on the F₁ was ~50% and similar proportions of pollen grains were inviable. We speculate that in the F₁ only male and female gametes containing the introgression are viable, and non-introgressed gametes are disabled. Further research is continuing on 38U16.

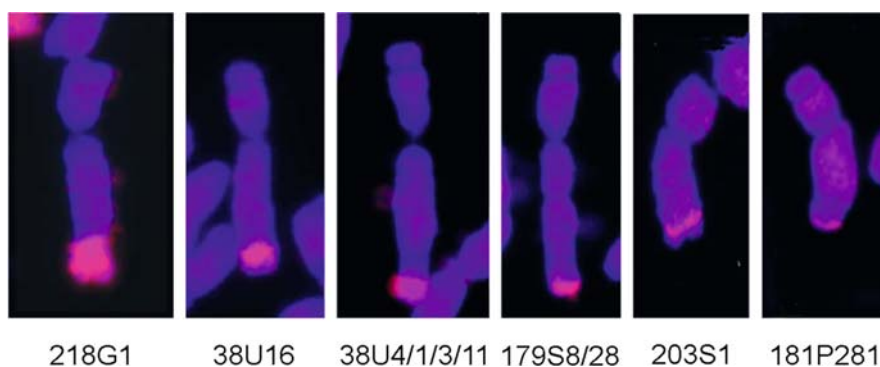


Fig. 5.5 Genomic *in situ* hybridisation of chromosome 5H derived from six *Hordeum vulgare* – *H. bulbosum* introgression lines. Different sized introgressions of *Hordeum bulbosum* chromatin (red signals) are seen on *H. vulgare* chromosome 5HL. Note that 38U16 and 203S1 are interstitial introgressions that lack *H. bulbosum* chromatin on the distal end of 5HL. The introgressions in 203S1, 218G1 and 38U16 encompass *Vrn-H1*, one of the loci that controls vernalisation requirement in barley (See Color Insert)

In summary, it is clear that over the past 10–15 years, molecular methods and cytogenetics have combined to form a powerful tool in determining species relationships within the genus and the results have aided the characterisation of introgression lines that have been developed for plant breeders. New marker technology and mapping strategies now becoming available and molecular markers will undoubtedly revolutionise the genetic analysis of performance traits and can be applied in marker-assisted selection schemes. The further integration of linkage maps with physical maps of DNA markers and genes will provide a better basis for basic research as well as for future barley breeding. In the future, improved sequencing technology and drastically reduced costs will result in a better understanding of the barley genome organisation and the forces and mechanisms involved in shaping them during evolution.

Acknowledgments The authors have been supported by the DFG, Germany (DFG HO 1779/9-1) and under an EU Marie Curie Research Fellowship (219313). R Pickering also thanks the Foundation for Research, Science and Technology (New Zealand) for financial support and his colleagues in Plant & Food Research, especially Paul Johnston, Viji Meiyalaghan and Stan Ebdon.

References

- Ananiev, E. (1992) Nuclear genome structure and organization. In: P.R. Shewry (Ed.), *Barley: Genetics, Biochemistry, Molecular Biology and Biotechnology*. CAB International, Wallingford, UK, pp. 133–150.
- Ashida, T., Nasuda, S., Sato, K. and Endo, T.R. (2007) Dissection of barley chromosome 5H in common wheat. *Genes Genet. Syst.* 82, 123–133.
- Banks, P.M., Larkin, P.J., Bariana, H.S., Lagudah, E.S., Appels, R., Waterhouse, P.M., Brettell, R.I.S., Chen, X., Xu, H.J., Xin, Z.Y., Qian, Y.T., Zhou, X.M., Cheng, Z.M. and Zhou, G.H. (1995) The use of cell culture for subchromosomal introgressions of barley yellow dwarf virus resistance from *Thinopyrum intermedium* to wheat. *Genome* 38, 395–405.
- Barakat, A., Carels, N. and Bernardi, G. (1997) The distribution of genes in the genomes of *Gramineae*. *Proc. Natl. Acad. Sci. USA* 94, 6857–6861.
- Baum, B.R. and Bailey, G. (1988) A taxonomic investigation of *Hordeum arizonicum* (*Poaceae: Triticeae*) with reference to related species. *Can. J. Bot.* 66, 1848–1855.
- Bayliss, M.W. and Riley, R. (1972) An analysis of temperature-dependent asynapsis in *Triticum aestivum*. *Genet. Res. Camb.* 20, 193–200.
- Becker, J. and Heun, M. (1995) Mapping of digested and undigested random amplified microsatellite polymorphisms in barley. *Genome* 38, 991–998.
- Bedbrook, J.R., Jones, J., Odell, M., Thompson, R.D. and Flavell, R.B. (1980) A molecular description of telomeric heterochromatin in *Secale* species. *Cell* 19, 545–560.
- Belostotsky, D.A. and Ananiev, E.V. (1990) Characterization of relic DNA from barley genome. *Theor. Appl. Genet.* 80, 374–380.
- Bennett, M.D., Finch, R.A. and Barclay, I.R. (1976) The time rate and mechanism of chromosome elimination in *Hordeum* hybrids. *Chromosoma* 54, 175–200.
- Bothmer, R. von, Flink, J., Jacobsen, N., Kotimäki, M. and Landström, T. (1983) Inter-specific hybridization with cultivated barley (*Hordeum vulgare* L.). *Hereditas* 99, 219–244.
- Bothmer, R. von, Flink, J. and Landström, T. (1987) Meiosis in *Hordeum* interspecific hybrids. II. Triploid hybrids. *Evol. Trend. Plant.* 1, 41–50.
- Bothmer, R. von, Jacobsen, N., Baden, C., Jørgensen, R.B. and Linde-Laursen, I. (1995) *An ecogeographical study of the genus Hordeum. (1995) 2nd edition. Systematic and Ecogeographical Studies on Crop Genepools* 7. IPGRI, Rome, Italy.
- Brandes, A., Röder, M.S. and Ganai, M.W. (1995) Barley telomeres are associated with two different types of satellite DNA sequences. *Chromosome Res.* 3, 315–320.
- Brown, S.E., Stephens, J.L., Lapitan, N.L.V. and Knudson, D.L. (1999) FISH landmarks for barley chromosomes (*Hordeum vulgare* L.). *Genome* 42, 274–281.
- Busch, W., Martin, R., Herrmann, R.G. and Hohmann, U. (1995) Repeated DNA sequences isolated by microdissection. I. Karyotyping of barley (*Hordeum vulgare* L.). *Genome* 38, 1082–1090.
- Busch, W., Herrmann, R.G., Houben, A. and Martin, R. (1996) Efficient preparation of plant metaphase spreads. *Plant Mol. Biol. Rep.* 14, 149–155.
- Cabrera, A., Friebe, B., Jiang, J. and Gill, B.S. (1995) Characterization of *Hordeum chilense* chromosomes by C-banding and in situ hybridization using highly repeated DNA probes. *Genome* 38, 435–442.
- Cho, S.H., Garvin, D.F. and Muehlbauer, G.J. (2006) Transcriptome analysis and physical mapping of barley genes in wheat-barley chromosome addition lines. *Genetics* 172, 1277–1285.

- Costa, J.M. and Singh, R.J. (2006) Chromosome mapping in barley (*Hordeum vulgare* L.). In: R.J. Singh and P.P. Jauhar (Eds.), *Genetic Resources, Chromosome Engineering, and Crop Improvement, Volume 2, Cereals*, CRC Press, Boca Raton, pp. 257–280.
- Cuadrado, A. and Jouve, N. (2007) The nonrandom distribution of long clusters of all possible classes of trinucleotide repeats in barley chromosomes. *Chromosome Res.* 15, 711–720.
- Dahleen, L.S. (1999) Tissue culture increases meiotic pairing of regenerants from barley \times Canada wild rye hybrids. *J. Hered.* 90, 265–269.
- de Bustos, A., Cuadrado, A., Soler, C. and Jouve, N. (1996) Physical mapping of repetitive DNA sequences and 5S and 18S-26S rDNA in five wild species of the genus *Hordeum*. *Chromosome Res.* 4, 491–499.
- Devaux, P. and Pickering, R. (2005) Haploids in the Improvement of Poaceae. In: C.E. Palmer, W.A. Keller and K.J. Kasha (Eds.), *Haploids in Crop Improvement II, Series: Biotechnology in Agriculture and Forestry, Vol 56*. Springer, Berlin, Heidelberg, pp. 215–242.
- Dolezel, J., Greilhuber, J., Lucretti, S., Meister, A., Lysak, M., Nardi, L. and Obermyer, R. (1998) Plant genome size estimation of flow cytometry: inter-laboratory comparison. *Ann. Bot. (Lond)* 82 (Supp. A), 17–26.
- Endo, T.R. (1988) Induction of chromosomal structural changes by a chromosome of *Aegilops cylindrica* L. in common wheat. *J. Hered.* 79, 366–370.
- Fedak, G. (1982) Effect of cultivar combination on meiosis in barley-wheat hybrids. *Can. J. Genet. Cytol.* 24, 575–582.
- Fernandez, J.A. and Jouve, N. (1988) The addition of *Hordeum chilense* chromosomes to *Triticum turgidum* conv. *durum*. Biochemical, karyological and morphological characterization. *Euphytica* 37, 247–259.
- Finch, R.A. and Bennett, M.D. (1982) The mechanism of somatic chromosome elimination in *Hordeum*. In: P.E. Brandham and M.D. Bennett (Eds.), *Kew Chromosome conference II*. London, pp. 146–153.
- Fukui, K. and Kakeda, K. (1990) Quantitative karyotyping of barley chromosomes by image analysis methods. *Genome* 33, 450–458.
- Fukui, K., Kamisugi, Y. and Sakai, F. (1994) Physical mapping of 5S rDNA loci by direct cloned biotinylated probes in barley chromosomes. *Genome* 37, 105–111.
- Fukuyama, T. and Hosoya, H. (1983) Genetic control and mechanism of chromosome elimination in the hybrids between *Hordeum bulbosum* (4 \times) and *Hordeum vulgare* (4 \times). *Jpn. J. Genet.* 58, 241–250.
- Gerlach, W.L. and Bedbrook, J.R. (1979) Cloning and characterization of ribosomal RNA genes from wheat and barley. *Nucleic Acids Res.* 7, 1869–1885.
- Gernand, D., Rutten, T., Varshney, A., Rubtsova, M., Prodanovic, S., Bruss, C., Kumlehn, J., Matzk, F. and Houben, A. (2005) Uniparental chromosome elimination at mitosis and interphase in wheat and pearl millet crosses involves micronucleus formation, progressive heterochromatinization, and DNA fragmentation. *Plant Cell* 17, 2431–2438.
- Gernand, D., Rutten, T., Pickering, R. and Houben, A. (2006) Elimination of chromosomes in *Hordeum vulgare* \times *H. bulbosum* crosses at mitosis and interphase involves micronucleus formation and progressive heterochromatinization. *Cytogenet. Genome Res.* 114, 169–174.
- Graner, A., Jahoor, A., Schondelmaier, J., Siedler, H., Pillen, K., Fischbeck, G., Wenzel, G. and Herrmann, R.G. (1991) Construction of an RFLP map of barley. *Theor. Appl. Genet.* 83, 250–256.
- Gupta, P.K. and Fedak, G. (1985a) Genetic control of meiotic chromosome pairing in polyploids in the genus *Hordeum*. *Can. J. Gen. Cyt.* 27, 515–530.
- Gupta, P.K. and Fedak, G. (1985b) Meiosis in seven intergeneric hybrids between *Hordeum* and *Secale*. *Z. Pflanzenzüchtg.* 95, 262–273.
- Harlan, J.R. and de Wet, J.M.J. (1971) Toward a rational classification of cultivated plants. *Taxon* 20, 509–517.

- Hart, G.E., Islam, A.K.M.R. and Shepherd, K.W. (1980) Use of isozymes as chromosome markers in the isolation and characterization of wheat-barley chromosome addition lines. *Genet. Res. Camb.* 311–325.
- Heddle, J.A. and Carrano, A.V. (1977) The DNA content of micronuclei induced in mouse bone marrow by gamma-irradiation: evidence that micronuclei arise from acentric chromosomal fragments. *Mutat. Res.* 44, 63–69.
- Ho, K.M. and Kasha, K.J. (1975) Genetic control of chromosome elimination during haploid formation in barley. *Genetics* 81, 263–275.
- Houben, A., Schroeder-Reiter, E., Nagaki, K., Nasuda, S., Wanner, G., Murata, M. and Endo, T.R. (2007) CENH3 interacts with the centromeric retrotransposon cereba and GC-rich satellites and locates to centromeric substructures in barley. *Chromosoma* 116, 275–283.
- Hudakova, S., Michalek, W., Presting, G.G., ten Hoopen, R., dos Santos, K., Jasencakova, Z. and Schubert, I. (2001) Sequence organization of barley centromeres. *Nucleic Acids Res.* 29, 5029–5035.
- Hudakova, S., Künzel, G., Endo, T.R. and Schubert, I. (2002) Barley chromosome arms longer than half of the spindle axis interfere with nuclear divisions. *Cytogenet. Genome Res.* 98, 101–107.
- Islam, A.K.M.R., Shepherd, K.W. and Sparrow, D.H.B. (1981) Isolation and characterization of euplasmic wheat-barley addition lines. *Heredity* 46, 161–174.
- Islam, A.K.M.R. (1980) Identification of wheat-barley addition lines with N-banding of chromosomes. *Chromosoma* 76, 365–373.
- Islam, A.K.M.R. (1983) Ditelosomic additions of barley chromosomes to wheat. In: S. Sakamoto (Ed.), *Proc 6th Int. Wheat Genetics Symp.* Kyoto University Press, Kyoto, Japan, pp. 233–238.
- Islam, A.K.M.R. and Shepherd, K.W. (1992a) Substituting ability of individual barley chromosomes for wheat chromosomes. *Plant. Breed.* 109, 141–150.
- Islam, A.K.M.R. and Shepherd, K.W. (1992b) Production of wheat-barley recombinant chromosomes through induced homoeologous pairing. *Theor. Appl. Genet.* 83, 489–494.
- Islam, A.K.M.R. and Shepherd, K.W. (2000) Isolation of a fertile wheat-barley addition line carrying the entire barley chromosome 1H. *Euphytica* 111, 145–149.
- Jensen, J. and Linde-Laursen, I. (1992) Statistical evaluation of length measurements on barley chromosomes with a proposal for a new nomenclature for symbols and positions of cytological markers. *Hereditas* 117, 51–59.
- Johnston, P.A. and Pickering, R. (2002) PCR detection of *Hordeum bulbosum* introgressions in an *H. vulgare* background using a retrotransposon-like sequence. *Theor. Appl. Genet.* 104, 720–726.
- Jørgensen, R.B. and Andersen, B. (1989) Karyotype analysis of regenerated plants from callus cultures of interspecific hybrids of cultivated barley (*Hordeum vulgare* L.). *Theor. Appl. Gen.* 77, 343–351.
- Kasha, K.J. and Kao, K.N. (1970) High frequency haploid production in barley (*Hordeum vulgare* L.). *Nature* 225, 874–876.
- Kasha, K.J. and Sadasivaiah, R.S. (1971) Genome relationships between *Hordeum vulgare* L. and *H. bulbosum* L. *Chromosoma* 35, 264–287.
- Kato, A., Albert, P., Vega, J. and Birchler, J. (2006) Sensitive fluorescence in situ hybridization signal detection in maize using directly labeled probes produced by high concentration DNA polymerase nick translation. *Biotech. Histochem.* 81, 71–78.
- Kim, N.S., Armstrong, K.C., Fedak, G., Ho, K. and Park, N.I. (2002) A microsatellite sequence from the rice blast fungus (*Magnaporthe grisea*) distinguishes between the centromeres of *Hordeum vulgare* and *H. bulbosum* in hybrid plants. *Genome* 45, 165–174.
- King, J., Armstead, I.P., Donnison, I.S., Harper, J.A., Roberts, L.A., Thomas, H., Ougham, H., Thomas, A., Huang, L. and King, I.P. (2007) Introgression mapping in grasses. *Chromosome Res.* 15, 105–113.

- Kleinhofs, A., Kilian, A., Maroof, M.A.S., Biyashev, R.M., Hayes, P., Chen, F.Q., Lapitan, N., Fenwick, A., Blake, T.K., Kanazin, V., Ananiev, E., Dahleen, L., Kudrna, D., Bollinger, J., Knapp, S.J., Liu, B., Sorrells, M., Heun, M., Franckowiak, J.D., Hoffman, D., Skadsen, R. and Steffenson, B.J. (1993) A molecular, isozyme and morphological map of the barley (*Hordeum vulgare*) genome. *Theor. Appl. Genet.* 86, 705–712.
- Koba, T., Handa, T. and Shimada, T. (1991) Efficient production of wheat-barley hybrids and preferential elimination of barley chromosomes. *Theor. Appl. Genet.* 81, 285–292.
- Kopyto, R., Crane, C.F. and Slepner, D.A. (1989) Effect of temperature on meiosis and fertility in *Festuca mairei* × *Festuca arundinacea* var. *glaucescens*. *Genome* 32, 708–718.
- Kumar, A. and Bennetzen, J.L. (1999) Plant retrotransposons. *Annu. Rev. Genet.* 33, 479–532.
- Künzel, G., Korzun, L. and Meister, A. (2000) Cytologically integrated physical restriction fragment length polymorphism maps for the barley genome based on translocation break-points. *Genetics* 154, 397–412.
- Künzel, G. and Waugh, R. (2002) Integration of microsatellite markers into the translocation-based physical RFLP map of barley chromosome 3H. *Theor. Appl. Genet.* 105, 660–665.
- Langdon, T., Seago, C., Mende, M., Leggett, M., Thomas, H., Forster, J.W., Jones, R.N. and Jenkins, G. (2000) Retrotransposon evolution in diverse plant genomes. *Genetics* 156, 313–325.
- Lange, W. (1971) Crosses between *Hordeum vulgare* L. and *H. bulbosum* L. II. Elimination of chromosomes in hybrid tissue. *Euphytica* 20, 181–194.
- Lange, W. and Jochemsen, G. (1976) Karyotypes, nucleoli, and amphiplasty in hybrids between *Hordeum vulgare* L. and *Hordeum bulbosum* L. *Genetica* 46, 217–233.
- Lapitan, N.L.V., Brown, S.E., Kennard, W., Stephens, J.L. and Knudson, D.L. (1997) FISH physical mapping with barley BAC clones. *Plant J.* 11, 149–156.
- Lehfer, H., Busch, W., Martin, R. and Herrmann, R.G. (1993) Localization of the B-hordein locus on barley chromosomes using fluorescence in situ hybridization. *Chromosoma* 102, 428–432.
- Leitch, I.J. and Heslop-Harrison, J.S. (1992) Physical mapping of the 18S-5.8S-26S rRNA genes in barley by in situ hybridization. *Genome* 35, 1013–1018.
- Leitch, I.J. and Heslop-Harrison, J.S. (1993) Physical mapping of 4 sites of 5S rDNA sequences and one site of the alpha-amylase-2 gene in barley (*Hordeum vulgare*). *Genome* 36, 517–523.
- Linde-Laursen, I. (1975) Giemsa C-banding of the chromosomes of 'Emir' barley. *Hereditas* 81, 285–289.
- Linde-Laursen, I., Bothmer, R. von and Jacobsen, N. (1990) Giemsa C-banded karyotypes of diploid and tetraploid *Hordeum bulbosum* (Poaceae). *Plant Syst. Evol.* 172, 141–150.
- Linde-Laursen, I., Bothmer, R. von and Jacobsen, N. (1992) Relationships in the genus *Hordeum*: Giemsa C-banded karyotypes. *Hereditas* 116, 111–116.
- Linde-Laursen, I., Heslop-Harrison, J.S., Shepherd, K.W. and Taketa, S. (1997) The barley genome and its relationship with the wheat genomes. A survey with an internationally agreed recommendation for barley chromosome nomenclature. *Hereditas* 126, 1–16.
- Malysheva, L., Sjakste, T., Matzk, F., Röder, M. and Ganai, M. (2003) Molecular cytogenetic analysis of wheat-barley hybrids using genomic in situ hybridization and barley microsatellite markers. *Genome* 46, 314–322.
- Manninen, O., Kalendar, R., Robinson, J. and Schulman, A.H. (2000) Application of BARE-1 retrotransposon markers to the mapping of a major resistance gene for net blotch in barley. *Mol. Gen. Genet.* 264, 325–334.
- Marthe, F. and Künzel, G. (1994) Localization of translocation breakpoints in somatic metaphase chromosomes of barley. *Theor. Appl. Genet.* 89, 240–248.
- Martin, R., Busch, W., Herrmann, R.G. and Wanner, G. (1996) Changes in chromosomal ultrastructure during the cell cycle. *Chromosome Res.* 4, 288–294.
- Martin, A. and Cabrera, A. (2005) Cytogenetics of *Hordeum chilense*: current status and considerations with references to breeding. *Cytogenet. Genome Res.* 109, 378–384.

- McIntyre, C.L., Pereira, S., Moran, L.B. and Appels, R. (1990) New *Secale cereale* (rye) DNA derivatives for the detection of rye chromosome segments in wheat. *Genome* 33, 635–640.
- Miller, T.E., Reader, S.M. and Chapman, V. (1982) The addition of *Hordeum chilense* chromosomes to wheat. In: C. Broertjes (Ed.), *Induced Variability in Plant Breeding*. Pudoc, Wageningen, pp. 79–81.
- Molnár-Lang, M., Linc, G., Friebe, B.R. and Sutka, J. (2000) Detection of wheat-barley translocations by genomic *in situ* hybridization in derivatives of hybrids multiplied *in vitro*. *Euphytica* 112, 117–123.
- Molnár, I., Linc, G., Dulai, S., Nagy, E.D. and Molnár-Lang, M. (2007) Ability of chromosome 4H to compensate for 4D in response to drought stress in a newly developed and identified wheat-barley 4H(4D) disomic substitution line. *Plant Breed.* 126, 369–374.
- Morrison, J.W. and Rajhathy, T. (1959) Cytogenetic studies in the genus *Hordeum*. III. Pairing in some interspecific and intergeneric hybrids. *Can. J. Gen. Cytol.* 1, 65–77.
- Morrison, J.W. (1959) Cytogenetic studies in the genus *Hordeum*. I. Chromosome morphology. *Can. J. Bot.* 37, 527–538.
- Murai, K., Koba, T. and Shimada, T. (1997) Effects of barley chromosome on heading characters in wheat-barley chromosome addition lines. *Euphytica* 96, 281–287.
- Nagy, E.D., Molnár-Lang, M., Linc, G., and Láng, L. (2002) Identification of wheat-barley translocations by sequential GISH and two-colour FISH in combination with the use of genetically mapped barley SSR markers. *Genome* 45, 1238–1247.
- Nasuda, S., Hudakova, S., Schubert, I., Houben, A. and Endo, T.R. (2005a) Stable barley chromosomes without centromeric repeats. *Proc. Natl. Acad. Sci. USA* 102, 9842–9847.
- Nasuda, S., Kikkawa, Y., Ashida, T., Islam, A.K.M.R., Sato, K. and Endo, T.R. (2005b) Chromosomal assignment and deletion mapping of barley EST markers. *Genes Genet. Syst.* 80, 357–366.
- Nilan, R. (1964) *The cytology and genetics of barley. Monographic supplement no. 3, Research studies*. Washington State University, pp. 277.
- Pedersen, C. and Linde-Laursen, I. (1994) Chromosomal locations of four minor rDNA loci and a marker microsatellite sequence in barley. *Chromosome Res.* 2, 65–71.
- Pedersen, C., Rasmussen, S.K. and Linde-Laursen, I. (1996) Genome and chromosome identification in cultivated barley and related species of the *Triticeae* (*Poaceae*) by *in situ* hybridization with the GAA-satellite sequence. *Genome* 39, 93–104.
- Pendinen, G.I. and Chernov, V.E. (1995) Characteristics of interaction of barley *Hordeum depressum* (Scribn. & Sm.) Ribd. (4×) with genomes of rye *Secale cereale* L. (2×) and barley *Hordeum vulgare* L. (2×) in meiosis of F_1 hybrids. *Russ. J. Genet.* 31, 323–329.
- Petersen, G. (1991) Intergeneric hybridization between *Hordeum* and *Secale*. *Hereditas* 114, 141–159.
- Pickering, R.A. (1984) The influence of genotype and environment on chromosome elimination in crosses between *Hordeum vulgare* L. × *Hordeum bulbosum* L. *Plant Sci. Lett.* 34, 153–164.
- Pickering, R.A. (1990) The influence of temperature on chromosome pairing in diploid and triploid hybrids between *Hordeum vulgare* L. and *H. bulbosum* L. *Hereditas* 113, 221–226.
- Pickering, R.A. (1992) Monosomic and double monosomic substitutions of *Hordeum bulbosum* L. chromosomes into *H. vulgare* L. *Theor. Appl. Genet.* 84, 466–472.
- Pickering, R.A. (1994) The chromosome stability of *Hordeum vulgare* L. – *H. bulbosum* L. substitution plants grown at two temperatures. *Hereditas* 121, 39–43.
- Pickering, R.A., Timmerman, G.M., Cromey, M.G. and Melz, G. (1994) Characterisation of progeny from backcrosses of triploid hybrids between *Hordeum vulgare* L. (2×) and *H. bulbosum* L. (4×) to *H. vulgare*. *Theor. Appl. Genet.* 88, 460–464.
- Pickering, R.A., Malyshev, S., Künzel, G., Johnston, P.A., Korzun, V., Menke, M. and Schubert, I. (2000) Locating introgressions of *Hordeum bulbosum* chromatin within the *H. vulgare* genome. *Theor. Appl. Genet.* 100, 27–31.
- Pickering, R., Niks, R.E., Johnston, P.A. and Butler R.C. (2004) Importance of Secondary and Tertiary Genepools in Barley Genetics and Breeding. II. Disease Resistance, Agronomic Performance and Quality. *Czech J. Genet. Plant Breed.* 40, 79–85.

- Pickering, R. and Johnston, P.A. (2005) Recent progress in barley improvement using wild species of *Hordeum*. Cytogenet. Genome Res. 109, 344–349.
- Pickering, R., Klatte, S. and Butler, R.C. (2005) Reduced chromosome association between the short arm of 5H homologues in *Hordeum vulgare* L. at metaphase I. Plant Breed. 124, 416–418.
- Pickering, R., Klatte, S. and Butler, R.C. (2006) Identification of all chromosome arms and their involvement in meiotic homoeologous associations at metaphase I in 2 *Hordeum vulgare* L. \times *Hordeum bulbosum* L. hybrids. Genome 49, 73–78.
- Presting, G.G., Malysheva, L., Fuchs, J. and Schubert, I. (1998) A TY3/GYPSY retrotransposon-like sequence localizes to the centromeric regions of cereal chromosomes. Plant J. 16, 721–728.
- Prieto, P., Ramírez, C., Cabrera, A., Ballesteros, J. and Martín, A. (2006) Development and cytogenetic characterisation of a double goat grass-barley chromosome substitution in tritordeum. Euphytica 147, 337–342.
- Qi, X. and Lindhout, P. (1997) Development of AFLP markers in barley. Mol. Gen. Genet. 254, 330–336.
- Rajhathy, T. and Symko, S. (1966) The synthesis of a species: *Hordeum arizonicum*. Can. J. Bot. 44, 1224–1228.
- Ramage, R. (1985) Cytogenetics. In: D.C. Rasmusson (Ed.), *Barley, Agronomy Monograph No. 26*. American Society of Agronomy, Madison, Wisconsin, pp. 127–154.
- Ramsay, L., Macaulay, M., Ivanissevich, S.D., MacLean, K., Cardle, L., Fuller, J., Edwards, K.J., Tuveson, S., Morgante, M., Massari, A., Maestri, E., Marmiroli, N., Sjakste, T., Ganai, M., Powell, W. and Waugh, R. (2000) A simple sequence repeat-based linkage map of barley. Genetics 156, 1997–2005.
- Randhawa, H.S., Dilbirligi, M., Sidhu, D., Erayman, M., Sandhu, D., Bondareva, S., Chao, S., Lazo, G.R., Anderson, O.D., Miftahudin, Gustafson, J.P., Echalié, B., Qi, L.L., Gill, B.S., Akhunov, E.D., Dvorak, J., Linkiewicz, A.M., Ratnasiri, A., Dubcovsky, J., Bermudez-Kandianis, C.E., Greene, R.A., Sorrells, M.E., Conley, E.J., Anderson, J.A., Peng, J.H., Lapitan, N.L.V., Hossain, K.G., Kalavacharla, V., Kianian, S.F., Pathan, M.S., Nguyen, H.T., Endo, T. R., Close, T.J., McGuire, P.E., Qualset, C.O. and Gill, K.S. (2004) Deletion mapping of homoeologous group 6-specific wheat expressed sequence tags. Genetics 168, 677–686.
- Rayburn, A.L. and Gill, B.S. (1986) Molecular identification of the D-genome chromosomes of wheat. J. Hered. 77, 253–255.
- Riley, R. and Chapman, V. (1958) Genetic control of cytologically diploid behaviour of hexaploid wheat. Nature (London) 182, 712–715.
- Rimpau, J., Smith, D.B. and Flavell, R.B. (1980) Sequence organization in barley and oats chromosomes revealed by interspecies DNA-DNA hybridization. Heredity 44, 131–149.
- Röder, M.S., Lapitan, N.L.V., Sorrells, M.E. and Tanksley, S.D. (1993) Genetic and physical mapping of barley telomeres. Mol. Gen. Genet. 238, 294–303.
- Rostoks, N., Park, Y.J., Ramakrishna, W., Ma, J., Druka, A., Shiloff, B.A., SanMiguel, P.J., Jiang, Z., Brueggeman, R., Sandhu, D., Gill, K., Bennetzen, J.L. and Kleinbarts, A. (2002) Genomic sequencing reveals gene content, genomic organization, and recombination relationships in barley. Funct. Integr. Genomics 2, 51–59.
- Ruge, B., Linz, A., Pickering, R., Proeseler, G., Greif, P. and Wehling, P. (2003) Mapping of *Rym14^{Hb}*, a gene introgressed from *Hordeum bulbosum* and conferring resistance to BaMMV and BaYMV in barley. Theor. Appl. Genet. 107, 965–971.
- Safar, J., Bartos, J., Janda, J., Bellec, A., Kubalakova, M., Valarik, M., Pateyron, S., Weiserova, J., Tuskova, R., Cihalikova, J., Vrana, J., Simkova, H., Faivre-Rampant, P., Sourdil, P., Caboche, M., Bernard, M., Dolezel, J. and Chalhou, B. (2004) Dissecting large and complex genomes: flow sorting and BAC cloning of individual chromosomes from bread wheat. Plant J. 39, 960–968.
- Schubert, I. and Oud, J.L. (1997) There is an upper limit of chromosome size for normal development of an organism. Cell 88, 515–520.

- Schubert, I., Shi, F., Fuchs, J. and Endo, T.R. (1998) An efficient screening for terminal deletions and translocations of barley chromosomes added to common wheat. *Plant J.* 14, 489–495.
- Schwarzacher, T. and Heslop-Harrison, J.S. (1991) In situ hybridization to plant telomeres using synthetic oligomers. *Genome* 34, 317–323.
- Serizawa, N., Nasuda, S., Shi, F., Endo, T.R., Prodanovic, S., Schubert, I. and Künzel, G. (2001) Deletion-based physical mapping of barley chromosome 7H. *Theor. Appl. Genet.* 103, 827–834.
- Shepherd, K.W. and Islam, A.K.M.R. (1992) Progress in the production of wheat-barley addition and recombination lines and their use in mapping the barley genome. In: P.R. Shewry (Ed), *Barley: Genetics, Biochemistry, Molecular Biology and Biotechnology*. CAB International, Wallingford, UK, pp. 99–114.
- Sherman, J.D., Smith, L.Y., Blake, T.K. and Talbert, L.E. (2001) Identification of barley genome segments introgressed into wheat using PCR markers. *Genome* 44, 38–44.
- Shi, F. and Endo, T.R. (1999) Genetic induction of structural changes in barley chromosomes added to common wheat by a gametocidal chromosome derived from *Aegilops cylindrica*. *Genes Genet. Syst.* 74, 49–54.
- Singh, R.J. and Tsuchiya, T. (1982) Identification and designation of telocentric chromosomes in barley by means of Giemsa N-banding technique. *Theor. Appl. Genet.* 64, 13–24.
- Singh, R.J. (2003) *Plant Cytogenetics*, 2nd edition. CRC Press, Boca Raton, USA.
- Sorokin, A., Marthe, F., Houben, A., Pich, U., Graner, A. and Künzel, G. (1994) Polymerase chain reaction mediated localization of RFLP clones to microisolated translocation chromosomes of barley. *Genome* 37, 550–555.
- Stein, N. (2007) Triticeae genomics: advances in sequence analysis of large genome cereal crops. *Chromosome Res.* 15, 21–31.
- Stephens, J.L., Brown, S.E., Lapitan, N.L.V. and Knudson, D.L. (2004) Physical mapping of barley genes using an ultrasensitive fluorescence in situ hybridization technique. *Genome* 47, 179–189.
- Subrahmanyam, N.C. and Kasha, K.J. (1973) Selective chromosomal elimination during haploid formation in barley following interspecific hybridization. *Chromosoma* 42, 111–125.
- Subrahmanyam, N.C. (1978) Meiosis in polyhaploid *Hordeum*: hemizygous ineffective control of diploid-like behaviour in a hexaploid? *Chromosoma* 66, 185–192.
- Suchankova, P., Kubalaková, M., Kovarova, P., Bartos, J., Cihalikova, J., Molnar-Lang, M., Endo, T.R. and Dolezel, J. (2006) Dissection of the nuclear genome of barley by chromosome flow sorting. *Theor. Appl. Genet.* 113, 651–659.
- Szakacs, E. and Molnár-Lang, M. (2007) Development and molecular cytogenetic identification of new winter wheat-winter barley ('Martonvasari 9 kr1'-'Igri') disomic addition lines. *Genome* 50, 43–50.
- Szigat, G. and Pohler, W. (1982) *Hordeum bulbosum* × *H. vulgare* hybrids and their backcrosses with cultivated barley. *Cereal Res. Commun.* 10, 73–78.
- Taketa, S., Ando, H., Takeda, K., Harrison, G.E. and Heslop-Harrison, J.S. (2000) The distribution, organization and evolution of two abundant and widespread repetitive DNA sequences in the genus *Hordeum*. *Theor. Appl. Genet.* 100, 169–176.
- Taketa, S., Linde-Laursen, I. and Künzel, G. (2003) Cytogenetic diversity. In: R. von Bothmer, T. van Hintum, H. Knüpfner and K. Sato (Eds.), *Diversity in Barley (Hordeum vulgare)*. *Developments in Plant Genetics and Breeding*, 7. Elsevier Science B.V., Amsterdam, The Netherlands, pp. 97–119.
- Taketa, S., Awayama, T., Ichii, M., Sunakawa, M., Kawahara, T. and Murai, K. (2005) Molecular cytogenetic identification of nullisomy 5B induced homoeologous recombination between wheat chromosome 5D and barley chromosome 5H. *Genome* 48, 115–124.
- Thomas, H.M. and Pickering, R.A. (1983) Chromosome elimination in *Hordeum vulgare* × *Hordeum bulbosum* hybrids. 2. Chromosome behavior in secondary hybrids. *Theor. Appl. Genet.* 66, 141–146.

- Timmerman, G.M., Pickering, R.A. and Melz, G. (1993) Characterization of *Hordeum vulgare* - *Hordeum bulbosum* chromosome substitution lines by restriction fragment length polymorphism analysis. *Genome* 36, 507–511.
- Tsuchiya, T. (1991) Chromosome mapping by means of aneuploid analysis in barley. In: P.K. Gupta and T. Tsuchiya (Eds.), *Chromosome Engineering in Plants: Genetics, Breeding, Evolution, Part A*. Elsevier Science B.V., Amsterdam, The Netherlands, pp. 361–384.
- Tsujimoto, H., Mukai, Y., Akagawa, K., Nagaki, K., Fujigaki, J., Yamamoto, M. and Sasakuma, T. (1997) Identification of individual barley chromosomes based on repetitive sequences: conservative distribution of Afa-family repetitive sequences on the chromosomes of barley and wheat. *Genes Genet. Syst.* 72, 303–309.
- Valarik, M., Bartos, J., Kovarova, P., Kubalaková, M., de Jong, J.H. and Dolezel, J. (2004) High-resolution FISH on super-stretched flow-sorted plant chromosomes. *Plant J.* 37, 940–950.
- Varshney, R.K., Grosse, I., Hahnel, U., Siefken, R., Prasad, M., Stein, N., Langridge, P., Altschmied, L. and Graner, A. (2006) Genetic mapping and BAC assignment of EST-derived SSR markers shows non-uniform distribution of genes in the barley genome. *Theor. Appl. Genet.* 113, 239–250.
- Walther, U., Rapke, H., Proeseler, G. and Szigat, G. (2000) *Hordeum bulbosum* – a new source of disease resistance – transfer of resistance to leaf rust and mosaic viruses from *H. bulbosum* into winter barley. *Plant Breed.* 119, 215–218.
- Wang, S., Lapitan, N.L.V. and Tsuchiya, T. (1991) Characterization of telomeres in *Hordeum vulgare* chromosomes by in situ hybridization. I. Normal diploid barley. *Jpn. J. Genet.* 66, 313–316.
- Werner, J.E., Endo, T.R. and Gill, B.S. (1992) Toward a cytogenetically based physical map of the wheat genome. *Proc. Natl. Acad. Sci. USA* 89, 11307–11311.
- Xu, J. and Snape, J.W. (1988) The cytology of hybrids between *Hordeum vulgare* and *H. bulbosum* revisited. *Genome* 30, 486–494.
- Xu, J., Procunier, J.D. and Kasha, K.J. (1990) Species-specific in situ hybridization of *Hordeum bulbosum* chromosomes. *Genome* 33, 628–634.
- Xu, J. and Kasha, K.J. (1992) Transfer of a dominant gene for powdery mildew resistance and DNA from *Hordeum bulbosum* into cultivated barley (*Hordeum vulgare*). *Theor. Appl. Genet.* 84, 771–777.
- Zhang, H.N., Nasuda, S. and Endo, T.R. (2000) Identification of AFLP markers on the satellite region of chromosome 1BS in wheat. *Genome* 43, 729–735.
- Zhang, L., Pickering, R.A. and Murray, B.G. (2001) A *Hordeum vulgare* × *H. bulbosum* tetraploid hybrid provides useful agronomic introgression lines for breeders. *N. Z. J. Crop Hortic. Sci.* 29, 239–246.

Chapter 6

Methods for Genetic Analysis in the *Triticeae*

Abraham Korol, David Mester, Zeev Frenkel, and Yefim Ronin

Abstract The objective of genetic analysis is to reveal genome structural and functional organization. One of the major tools developed at early stages of genetics was genetic mapping. Genetic maps are a very important tool in evolutionary genomics and numerous practical applications like breeding, medical genetics, and gene cloning. An important usage of multilocus maps is genetic dissection of quantitative traits, or mapping quantitative trait loci (QTL). Fine QTL mapping is a prerequisite for efficient marker-assisted selection and map-based cloning. However, the fine mapping challenge, especially if the target is a gene of weak or moderate effect, requires large sample sizes and dense maps. New array-based technologies (SNP and tilling arrays) partially solve this problem but at a very high project-wise genotyping cost. This is why despite some technical obstacles, genetic analysis based on selective genotyping and selective DNA pooling becomes very popular, especially in human genetics. In this chapter we consider methods for building genetic maps (Section 6.1), various versions of “multiple” approach for QTL mapping (Section 6.2), and a new cost-effective method for genetic mapping based on selective DNA pooling (Section 6.3). Whenever possible, the examples are based on *Triticeae* species.

6.1 Construction of High Quality Dense Genetic Maps

New high throughput DNA technologies resulted in a disproportion between the high number of scored markers and relatively small size of mapping populations. Correspondingly, the number of scored markers may by orders of magnitude exceed the recombination resolution achievable for given population size. Hence, only a small part of markers can be genuinely ordered in the map. The question is how to choose the most informative markers for building such a “skeleton” or “framework” map. In this section, we discuss methods and

A. Korol (✉)

Institute of Evolution, Faculty of Science, University of Haifa, Haifa, Israel
e-mail: korol@research.haifa.ac.il

algorithms to address multilocus mapping problems including: (i) tools for multilocus ordering; (ii) map verification procedures; (iii) complicated genetic situations caused by “pseudo-linkage” and negative interference or meiotic gene conversion; (iv) algorithm to increase stability of multilocus ordering; and (v) principles of consensus mapping.

6.1.1 Multilocus Ordering

A primary difficulty in ordering genetic loci using linkage analysis is the large number of possible orders: for n loci on a chromosome, $n!/2$ distinct orders should be evaluated. In real problems, n may vary from dozens to 200–500 markers and more. One of the simplest possibilities in addressing the problem is to recover the linear marker order from known matrix d_{ij} of pairwise marker distances. Clearly, even for $n \sim 30$, it would not be feasible to evaluate all $n!/2$ possible orders using two-point linkage data. This is why multilocus ordering is considered as a computationally hard combinatorial optimization problem. We developed a highly efficient algorithm of multilocus ordering using two-locus linkage data that employs the Evolutionary Optimization Strategy (ES) (Mester et al. 2003a). It is based on the similarity between the mapping problem and the well known traveling salesperson problem (TSP) (Week and Lange 1987). ES belongs to a group of heuristic algorithms of discrete optimization mimicking natural population processes. With this approach, each multilocus order is considered as a “genotype” having “selective value” (or “fitness”) associated with map length: the shorter the map the higher is the fitness. Then a population of genotypes is allowed to evolve under various forces (mutation, recombination, reproduction, and selection) until the “best” genotype (i.e., order with the shortest map length) is obtained (Mester et al. 2003a).

The minimal-map-length criterion was employed by many other authors. However, using marker orders rather than marker positions (in cM) as the main map characteristic, is the core feature of our approach. Consequently, the stability of multilocus ordering rather than the sizes of confidence intervals of marker cM positions (Liu 1998; Jansen et al. 2001) as a criterion of the map quality, is central to our method. Marker order is a much more objective indicator for comparison multipoint maps than map positions. Indeed, even with strict constancy of gene order within species, recombination rates (hence map positions) may widely fluctuate among experiments due to sampling variation, dependence on ecological conditions, sex, genotype, and age (Korol et al. 1994). Therefore, genetic mapping of any target trait, either qualitative or quantitative, through determining the marker brackets, is less dependent on such fluctuations and more comparable across studies than a related effort based on map position in centiMorgans. This is especially clear when fine mapping serves a starting point for map-based cloning. Then, what is really important is the information about close markers rather than the map position in cM.


With this approach, testing the stability of the resulting map can be conducted by iterative re-sampling from the initial dataset followed by repeated rebuilding of the map and evaluating the stability of marker neighborhoods (Mester et al. 2003a). Such computing intensive strategy requires tools for fast optimization. To improve the efficiency of our multilocus ordering algorithms, we developed a new metaheuristic approach, referred to as Guided Evolution Strategy, that combines the strengths of Guided Local Search (Voudouris 1997) and Evolution Strategies (Mester et al. 2003a) in one iterative two-stage procedure (Mester et al. 2004). The last versions of this algorithm proved efficient with up to 3000 markers per chromosome.

In many practical cases, high density mapping is associated with another difficult problem: a disproportion between a high number of scored markers (k) and small sample size (N). The number of markers may by orders of magnitude exceed the mapping resolution for the given N . With a relatively high k/N ratio, we may observe groups of markers that have identical profiles along the population. Such groups should be replaced by their “delegates” that could be reliably ordered, comprising a “skeleton” map (Mester et al. 2003a). With $N \sim 100$, the minimum distance that can be resolved in the map should be ≥ 1 cM; thus for a chromosome with $k \sim 1000$ markers the map length should be ≥ 1000 cM, which is unrealistic for most species. How can be explained the appearance of 1000 cM maps? We believe that the root is scoring errors that “convert” clusters of *identical* (for given population size) markers into sets of non-identical *closely linked* markers. This is complemented by the naive assumption that all non-identical markers are resolvable by recombination and can be unequivocally ordered in a linear order. Although the relative order of markers absolutely linked to their delegates cannot be resolved in the genetic map for the given size of the mapping population, their presence may be helpful in physical mapping for anchoring of the contigs to the genetic map. Complementary, the higher resolution in physical mapping may allow relative positioning of such markers unachievable in genetic mapping with a small or modest size of the mapping population.

6.1.2 Map Verification Procedures

The objective of the verification procedure is detecting regions with unstable neighborhoods. This can be achieved by a *jackknife* procedure (Efron 1979): repeated re-sampling of the initial data set, e.g., using each time a certain proportion (say, 80%) of randomly chosen genotypes of the mapping population. For each such sub-sample one may build the best multilocus map, and compare the resulting maps to the map of the full sample. Then, the identification of unstable regions can be conducted based on the frequency distribution of the right-side and left-side neighbors for each marker. The higher the deviation from 1 (i.e., from “diagonal” pattern) the less certain is local order (Fig. 6.1). Clearly, the unstable neighborhoods result from fluctuations of

Name	Missing	Segregation	Index	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Xgwm219(1)	0	0.8:1.6:10	1	1	.2																					
Xgwm108a(2)	1	2.5:1.0(d4)	2	1	.8																					
Xgwm626(3)	0	0.9:1.7:10	3	.2	.8	1																				
P57M51f(4)	0	2.7:1.0(d4)	4			1	.7	.7	.3																	
P56M53u(5)	2	2.4:1.0(d4)	5				.7	.8	.3	.1			.1													
P56M60h(6)	3	3.0:1.0(d4)	6					.8	.8	.1	.1			.1	.1											
Xgwm582a(7)	2	0.9:1.8:1.0	7					.3	.3	.8	.6															
Xgwm58(8)	0	0.8:1.7:1.0	8					.1	.1	.6	.8	.1					.2	.1								
Xgwm88a(9)	0	0.9:1.9:1.0	9					.1	.8	.8	.8				.3											
Xgwm88c(10)	0	2.8:1.0(d4)	10						.1	.8	.8		.8			.3										
Xgwm107c(11)	0	0.9:2.1:1.0	11				.1				.8		.9	.2												
P56M50r(12)	3	3.3:1.0(d4)	12				.1					.9	.9	.2			.1									
P57M53d(13)	0	3.1:1.0(d4)	13							.3	.2	.9		.6												
Xgwm361c(14)	41	1.0:1.8:0.9	14					.1			.3				.6	1										
Xgwm193(15)	0	0.8:2.0:1.0	15						.2						1		.6	.2								
Xgwm508(16)	0	0.8:1.8:1.0	16							.1				.1			.6	1	.2							
Xgwm132a(17)	3	0.9:1.8:1.0	17														.2	1	.8							
P55M60q(18)	0	2.8:1.0(d4)	18															.2	.8		.9		.1			
P56M53b(19)	3	3.7:1.0(d4)	19																	.9		1	.1			
P55M60l(20)	0	3.4:1.0(d4)	20																			1	.9	.1		
Xgwm613(21)	0	1.0:2.1:1.0	21																	.1	.1	.9		.9		
P57M49a(22)	1	3.4:1.0(d4)	22																				.1	.9		1
P57M53o(23)	0	3.3:1.0(d4)	23																						1	



Name	Missing	Segregation	Index	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Xgwm219(1)	0	0.8:1.6:10	1	1	1															
Xgwm108a(2)	1	2.5:1.0(d4)	2	1	1	1														
Xgwm626(3)	0	0.9:1.7:10	3	1	1	1														
Xgwm582a(7)	2	0.9:1.8:1.0	4		1	1	1													
Xgwm58(8)	0	0.8:1.7:1.0	5			1	1	1												
Xgwm88a(9)	0	0.9:1.9:1.0	6				1	.9	.1											
P57M53d(13)	0	3.1:1.0(d4)	7					.9	1	.1		.1								
Xgwm107c(11)	0	0.9:2.1:1.0	8						1	1										
P56M50r(12)	3	3.3:1.0(d4)	9					.1	1	.9										
Xgwm193(15)	0	0.8:2.0:1.0	10						.1	.9		1								
Xgwm508(16)	0	0.8:1.8:1.0	11									1	1							
Xgwm132a(17)	3	0.9:1.8:1.0	12										1	1						
P55M60q(18)	0	2.8:1.0(d4)	13											1	1					
P56M53b(19)	3	3.7:1.0(d4)	14												1	.9	.1			
Xgwm613(21)	0	1.0:2.1:1.0	15													.9	1	.1		
P57M49a(22)	1	3.4:1.0(d4)	16														.1	1	.9	
P57M53o(23)	0	3.3:1.0(d4)	17														.1	.9		

Fig. 6.1 Re-sampling based procedure for detecting unstable neighborhoods followed by map stabilization by removing problematic markers (example on wheat chromosome 2B using data of Peng et al. 2000)

estimates of recombination rates across the repeated samples; the range of fluctuations depends on the sample size and the proportion of genotypes sampled at each jackknife run. In fact, this analysis is a modeling tool to quantify the diversity of map versions for the treated chromosome representing the sampling (stochastic) nature of the map. The results of such evaluation can be visualized to facilitate decision making about candidate problematic markers that should be removed from the map with following re-building of the map (see Section 6.1.4).

6.1.3 Complication due to ‘Pseudo-Linkage’ and Negative Interference

Correct multilocus genetic mapping includes two basic objectives: (i) Markers that belong to non-homologous chromosomes should not be assigned to the

same linkage group; and (ii) Markers from the same chromosome should be placed on the genetic map in the same order as the corresponding DNA sequences reside in the chromosome. Various factors may embarrass these goals, including deviation from independent segregation of non-syntenic loci (i.e., those that belong to non-homologous chromosomes) and excessive double recombinants.

In situations with significant deviations of the recombination rates, r , between non-syntenic markers from the expected 50%, the problem of marker clustering into linkage groups cannot be solved by an arbitrary choice of a certain threshold value of recombination or LOD. Indeed, in experiments with the foregoing characteristics, the recombination values between groups of markers from different chromosomes may be smaller than between adjacent markers within a chromosome. This phenomenon, referred to as *quasi-linkage* (or *pseudo-linkage*) can result from a combination of biological reasons and scoring errors. Pseudo-linkage was observed in quite a few of animal and plant species including cereals (see: Korol et al. 1994; Peng et al. 2000; and references therein). Clearly, quasi-linkage should be discriminated from another basic factor as interchromosomal translocations, which is not uncommon in *Triticeae* species (e.g., Liu et al. 1992; Joppa et al. 1995). Fifty years ago, the departure from random segregation of non-homologous chromosomes, observed in mouse crosses, was explained by a tendency of segregating chromosomes of the same origin (i.e., maternal or paternal) to co-migrate to the same pole during meiosis (the ‘affinity’ hypothesis) (Michie 1953). Consequently, gametes with parental combinations of non-homologous chromosomes should appear at a higher frequency than recombinant gametes. Beside non-random assortment, quasi-linkage can be caused by differential viability. Evidence of this type was found in *Drosophila* (Dobzhansky et al. 1965). Selective differences between parental and recombinant combinations may occur at the levels of gametes, zygotes, embryos, and adults. The effect can also be caused by selective fertilization during syngamy (Korol et al. 1994).

Numerous deviations from random segregation of non-homologues were detected in our F_2 mapping population *T. durum* \times *T. dicoccoides* (Peng et al. 2000). These included excess or deficit of parental genotypes, with the following highest deviating r values: $r(1A,6A)=0.78$, $r(2A,2B)=0.77$, $r(2A,7B)=0.27$, $r(5A,7A)=0.72$, and $r(3B,5B)=0.73$. Testing the significance of the observed deviations is complicated by the problem of multiple comparisons (91 chromosome-pair combinations, each with multiple markers per chromosome). Having in mind these complications, in evaluating the genome-wise significance of the observed numerous manifestations of quasi-linkage, we employed a permutation test. The result was that quasi-linkage for this population could be declared “significant” with $p<0.02$. Another example is based on our re-analysis of data on barley dihaploid population Steptoe \times Morex taken from the GrainGenes repository (<http://wheat.pw.usda.gov/>). Abundant coupling phase quasi-linkage associations were found for this population. Out of 21 chromosome pairs, 15 showed significant deviation from the expected 50% at $p<0.005$ – 0.01 .

After false discovery rate correction, 11 were found to be significant at $FDR = 1\%$: $r(1H,3H) = 0.37$, $r(1H,4H) = 0.27$, $r(1H,6H) = 0.34$, $r(2H,3H) = 0.22$, $r(2H,6H) = 0.38$, $r(3H,4H) = 0.38$, $r(3H,5H) = 0.36$, $r(3H,6H) = 0.36$, $r(4H,6H) = 0.35$, $r(5H,6H) = 0.36$, and $r(6H,7H) = 0.36$.

Mapping algorithms tend to ignore pseudo-linkage. Consequently, some non-syntenic loci may appear in the same "linkage group" resulting in contradictions between mapping results for different populations and between genetic and physical maps. Thus, up to 12% (!) of cattle markers were assigned to wrong chromosomes and contradicted to physical maps (H. Lewin, personal communication). In fish genetics, pseudo-linkage is also a known phenomenon (see Korol et al. 2007a). To address this problem, we suggested a modified scheme of building linkage groups, where clustering is conducted concurrently with multitilocus ordering and verification of the order and removing unreliable markers (Mester et al. 2003a). Instead of using a single threshold r value (or LOD value), we employ a series of increasing r (or decreasing LOD) thresholds. At the first, the most stringent threshold, we have a minimum danger of mixing markers from different chromosomes into one linkage group, but the cost is high number of linkage groups. By relaxing the stringency at the next steps, we allow end-to-end merging of the ordered linkage groups, excluding those that display strong affinity to each other by their interior parts.

Positive crossover interference, i.e., a reduced frequency of double crossovers compared to that expected with the assumption of independence, is a characteristic of meiotic organisms, with only a very few exceptions (reviewed in: Korol et al. 1994). Cases are known of higher than expected frequency of double crossovers in adjacent segments of small genetic but large physical length. In *Drosophila melanogaster*, a significant excess of multiple exchanges was found within a segment 4 cM long spanning the centromere and accounting for 25% of the cytological length of chromosome 3 (Sinclair 1975). Similar results have been obtained in other *Drosophila* studies with autosomes (see Korol et al. 1994). Significant negative crossover interference was found in barley (SØgaard 1977; Esch and Weber 2002). Negative crossover interference was found in our *T. durum* \times *T. dicoccoides* hybrid (Peng et al. 2000). As a rule, islands of negative interference were located in proximal regions, either spanning the centromere (chromosomes 1A, 3B, 4A, 6B, 7A, 7B), or proximal to it. The revealed trends were confirmed by using larger intervals from one or both sides of the node point of any considered interval pair and by variation of the position of the node. Islands of negative interference were also detected in *Ae. tauschii* (TA1691 \times TA1704) hybrid, with a tendency for a higher level of negative interference in regions that were proximal to or spanned the centromere (especially in chromosomes 2D, 3D, 4D, and 5D) (Boyko et al. 2002). We found a similar pattern when re-analyzing the data on rye (Korzun et al. 2001). Remarkably, like in wheat, the islands of negative interference in rye were mainly associated with centromeric regions (Fig. 6.2). These findings corroborate the hypothesis of Denell and Keppy (1979) that negative chromosome interference could be a characteristic of genomic regions exhibiting very low

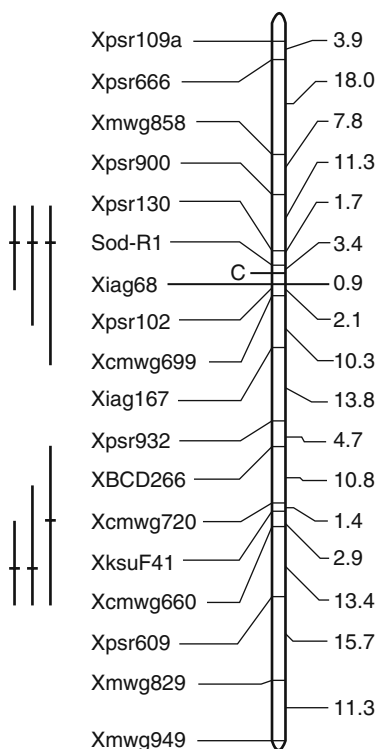


Fig. 6.2 Two islands of negative interference in chromosome 2 of rye. The relative independence of the effect is confirmed by changing the “node” marker of the analyzed adjacent interval (marked by a bold point) and by varying the upper or lower flank marker. Note that one of the two islands is in the vicinity of the centromere (marked by symbol C). The results are based on data of F_2 ($P87 \times P105$) mapping population (Korzun et al. 2001)

recombination rate per unit physical length. It is worth mentioning a few recent studies of relative contribution of crossing-over vs. gene conversion based on population sequence data on several organisms, including wild barley (Morrell et al. 2006) and *Arabidopsis* (Plagnol et al. 2006). It appeared that about a half of recombination events are represented by gene conversions. No matter what is the real mechanism of negative interference, it can play an important role in gene transfer in introgressive hybridization (hence in breeding) as well as a factor affecting fidelity of multilocus mapping.

Like negative interference, scoring errors during population genotyping for marker loci may also lead to excessive recombinants (in such cases, false recombinants). The stability of the negative-interference pattern to sliding scanning along flanking markers for a fixed node marker and vice versa, makes implausible the assumption of reading errors as an explanation for the aforementioned cases. However, reading errors seem to be a non-negligible factor in some mapping projects. Both, true negative interference and false

double recombinants may cause map expansion due to violation of the simple principle “the entire is larger than its parts” (Mester et al. 2003a). Therefore, it is a usual practice (not rational, to our mind) that after multilocus ordering, double recombinants are replaced by non-recombinants, resulting in “shorter” maps. This violation may also increase the chances of wrong multilocus ordering, although it appeared to have a minor effect on ordering based on our algorithm (see Mester et al. 2003a).

6.1.4 Increasing the Stability of Multilocus Maps

After revealing the regions of map instability based on re-sampling procedure, we need to make a decision about the markers responsible for the local instability. Clearly, our choice should depend on information quality of the markers (anchor markers or genes versus other markers; co-dominant versus dominant; fully scored or with many missing data; etc.). Having these criteria in mind, we can check the effect of removing any of the candidates using trial-and-error approach: after removing of a candidate marker, we can re-build the map, again test its stability based on the jackknife resampling, etc. As a result, we will get a stabilized skeleton map (see Fig. 6.1). The core procedure in our approach includes the following stages (see Fig. 6.3): (i) initial clustering with a stringent threshold r or LOD; (ii) multilocus ordering; (iii) bounding together closely linked markers followed by selection of “delegates” markers (bin markers) and replacing the groups of tightly linked markers by their “delegates”; (iv) repeated ordering and verification of the reduced linkage groups (LGs) to detect regions of map instability; (v) removing the markers causing unstable neighborhoods and violating monotonic change of recombination, followed by repeated ordering to get the skeleton map; (vi) stepwise relaxing the threshold r or LOD and merging end-to-end LGs followed by repeated application of steps (ii)–(v) (ideally, until their number fits the species chromosome number); (vii) attaching the removed markers to their best intervals on the skeleton map.

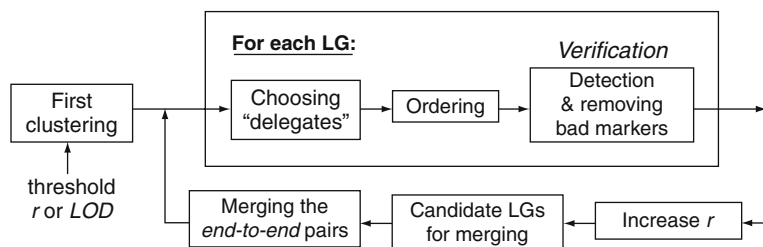


Fig. 6.3 The proposed scheme of multilocus mapping based on stepwise clustering of markers into linkage groups coordinated with multilocus ordering

To illustrate the advantages of the proposed methodology on cereal data, we need some criteria of comparisons. These include:

1. *Map length and number of markers presented in the resulting stable map:* Clearly, even with correctly recovered marker order, the evaluated map length may be higher than the true one, due certain amount of scoring errors.
2. *Controlling monotony:* Some errors may generate situations with violation of the principle “the entity is larger than its parts”. Normally, for three markers, ordered as *a-b-c*, one would expect: $r_{ac} > r_{ab}$ & $r_{ac} > r_{bc}$.
3. *Stability of local neighborhoods:* With reading errors, the ordering may be sensitive to variation among jackknife runs. As a stability measure of the constructed map we use $\sigma^2 = (1/k) \sum \sigma_i^2$, where i denotes the i th position in the multilocus order calculated on the entire sample, k is the number of markers in the order, σ_i^2 is variance of i th marker’s neighborhood, $\sigma_i^2 = 0.5 \sum p_{ij}(i-j)^2$, and p_{ij} is the proportion of jackknife runs, where markers i and j were adjacent. Stable ordering corresponds to $\sigma = 1$.
4. *Concordance of segregation distortions:* Upon correct ordering, one would expect a correspondence between segregation ratios of neighbor markers i and $i+1$. Correct order should give smaller values of the following criterion S_i than wrong order. If the frequency of one of the two classes at a marker locus in RIL (or dihaploid) population or one of the homozygotes in F_2 is p_i , then the normalized change of segregation from i to $i+1$ can be calculated as $S_i = 100 \cdot d_i / r_{i,i+1}$, with $d_i = |p_{2i} - p_{2i+1}|$, where p_{2i} and p_{2i+1} are frequencies of the second marker class for loci i and $i+1$ normalized by p_{1i} and p_{1i+1} , and $r_{i,i+1}$ is the distance (in cM) between the markers. For F_2 , $d_i = \max\{|p_{2i} - p_{2i+1}|, |p_{3i} - p_{3i+1}|\}$, where p_{2i} and p_{3i} are the normalized frequencies of the heterozygote and the second homozygote.

An empirical assessment of the proposed analytical framework, in comparison with other procedures, can be conducted using simulated and real data. Validation of the basic properties of our algorithms was provided earlier on simulated data (Mester et al. 2003a,b). Examples in Table 6.1 based on public domain cereal mapping data show that considerable improvement can be achieved in some cases with respect to map stability and map length. Based on our previous intensive testing (Mester et al. 2003a,b), we believe that the observed improvement of real data solutions reflects a better correspondence of the revised maps, compared to the initial ones, to the genuine chromosomal order of loci.

6.1.5 Building Consensus Maps

Numerous mapping projects conducted on cereals have generated an abundance of mapping data. Consequently, many multilocus maps were constructed using diverse mapping populations and marker sets for the same plant. The

Table 6.1 Examples of comparing the quality characteristics of the revised multilocus maps with those of the original maps

				L, cM				
Plant	Chr	σ		# markers		$D(ac)/D(ab \text{ or } bc)$		
Oat	MN3	MM	MP	MM	MP	MM		
				1.30	1.00	84	56	$D(3,8)/D(3,4) = 0.82$
						10	5	$D(8,10)/D(9,10) = 0.77$
Barley	6H	JM	MP	JM	MP	JM		
				1.39	1.04	175	157	$D(27,31)/D(30,31) = 0.85$
						37	35	
Wheat	1B	JM	MP	JM	MP	JM		
				4.05	1.04	445	138	$D(15,18)/D(15,16) = 0.55$
						72	26	$D(15,18)/D(16,17) = 0.71$
Rye	7R	JM	MP	JM	MP	JM		
				1.27	1.04	126	119	$D(8,10)/D(8,9) = 0.77$
						17	17	$D(8,10)/D(9,10) = 0.90$

Notes: σ – score of order instability (averaged for the chromosome); L – length of the map; $D(ac)/D(ab \text{ or } bc)$ – ratio of the map distance (cM) between flanking markers of a segment to the length of one of its parts ($D(ac)/D(ab \text{ or } bc) < 1$ indicates wrong marker scoring, wrong place of the internal marker, or negative interference); MM, JM, MP – the map was constructed with Mapmaker, Joinmap, or MultiPoint. Coding marker names: **oat** – 3=p38m35n2, 4=p35m68m6, 5=cdo608x, 8=cdo1467; **barley** – 27=ksuA3D, 30=MWG684B, 31=Nar7; **wheat** – 15=XksuE19, 16=Xbcd340, 17=Xrz166, 18=Xbcd1124; **rye** – 8=Xiag89, 9=Xpsr303, 10=Prx10.

quality of maps varied broadly between populations, marker sets, and used software. As one would expect, there might be some inconsistencies between different versions of the maps for the same organism. This proved to be the case for many organisms calling for efforts to integrate the mapping information and generate consensus maps (Klein et al. 2000; Menotti-Raymond et al. 2003). The problem of consensus mapping is even more challenging compared to multilocus mapping based on one data set, due to additional complications: differences in recombination rate and exchange distribution along chromosomes; variations in dominance of the employed markers; different subsets of markers used by different labs, hence the need for handling arbitrary patterns of shared sets of markers.

Usually, “merging” genetic maps is based on “giving credit” to the available maps and looking for shared orders of maximum length among the maps resulted from different projects (e.g., Yap et al. 2003; Emrich et al. 2004; Givry et al. 2005). A paradoxical situation may arise with this approach when with increasing number of mapping populations (and increasing amount of mapping information) the size of consensus orders will fall down.

An alternative developed in our lab is building de novo consensus maps based on joint re-analysis of the initial data sets rather than merging the previously constructed maps. Several problems can be formulated in terms of multilocus consensus genetic mapping (MCGM) (Mester et al. 2005). The main aspect of MCGM approach is requirement of identical order of shared markers

for any set and subset of mapping populations. Various formulations of MCGM problems can be considered:

- (A) *Multilocus genetic maps with “dominance” complication*: The problem appears at the building of genetic maps using F_2 or F_3 data with dominant markers in repulsion phase. In such situations we split the marker set into two groups, each with dominant markers in coupling phase only and the shared codominant markers (Peng et al. 2000). Multilocus maps are then ordered for the two sets with a requirement that shared (codominant) markers should have equal order (Mester et al. 2003b).
- (B) *Multilocus genetic maps with sex-dependent recombination*: These maps are built on the basis of male and female recombination rates represented as sex-specific matrices. Sex-specificity of the “distance” matrix may force an optimization algorithm to produce different marker orders (maps). Thus, our goal is to find the optimal solution with the restriction of the *same order* in two maps.
- (C) *Multilocus ordering in building consensus maps*: Such maps should be constructed based on re-analysis of raw recombination data generated in different genomic centers on different mapping populations.

A special feature of our approach to MCGM in case (C) is that we conduct multilocus mapping with a request of consensus order for all shared markers by analysis of initial data. This is in contrast to a more usual approach (e.g., Yap et al. 2003; Emrich et al. 2004; Givry et al. 2001), where the analysis is conducted on previously constructed maps by detecting and removing conflicting markers. Our formulations of MCGM can be considered in terms of connected (synchronized) travelling salesperson problem (Mester et al. 2005). To address this problem we also employ Evolution Strategy approach for generating consensus orders for shared markers. Searching the optimal order is based on our previously published highly efficient Guided Evolution Strategy algorithm (Mester et al. 2004). GES was modified for working with shared markers. Furthermore, a generated order of shared markers is used for each of the synchronously treated population as “anchor”. Such anchor-based ordering of non-shared markers implies solving TSP with difficult constraints. Several authors employed TSP approach for genetic and physical mapping (e.g., Weeks and Lange 1987; Falk 1992; Schiex and Gaspin 1997; Ben-Dor et al. 2000), but heuristics proposed in these publications cannot deal with constraint TSP. Optimal consensus order of shared markers is defined by minimal total length of maps of the chromosome. This criterion can be modified to account of variation in quality of the original datasets (population size, standard deviation after jackknife verification, etc.).

We provide some illustrations for the presented formulations (A) and (C) using simulated data, so that one can get an idea what was the problem and how it is solved. The first example is on map construction based on a mixture of dominant and codominant markers (see Mester et al. 2003b). We split the data on two subsets and then order the two

with a constraint that shared (codominant) markers appear in identical order in the maps. In the simulated example on F_2 , 10 out of 30 markers were codominant (underlined in the illustration below). Separate non-synchronized ordering resulted in wrong orders (NS-orders) highlighted in grey (the markers were named by sequential numbers as they should appear in the correct map). On the contrary, synchronized ordering resulted in recovery of the correct order (S-orders).

Set 1	NSO	1 2 3 4 5 6 7 8 10 14 15 29 30 28 27 26 25 22 21 20
	SO	1 2 3 4 5 6 7 8 10 14 15 20 21 22 25 26 27 28 29 30
Set 2	NSO	17 16 14 13 12 2 5 19 20 21 22 23 24 25 27 28 30
	SO	2 5 12 13 14 16 17 19 20 21 22 23 24 25 27 28 30

Next example represents joint ordering of five mapping populations (formulation C). As before, we underline the shared markers; wrong orders are marked in grey.

Pop1	NSO	9 8 7 6 5 4 3 2 1 29 30 28 27 26 25 24 23 22 21 20 19 18 17 16 15 14
	SO	1 2 3 4 5 6 7 8 9 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
Pop2	NSO	1 2 3 4 5 6 7 8 10 9 11 12 13 14 15 16 17 18 19 20 21 30 29 28
	SO	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 28 29 30
Pop3	NSO	20 19 18 17 16 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1 25 26 27 28 29 30
	SO	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 25 26 27 28 29 30
Pop4	NSO	18 19 20 21 22 23 24 25 26 27 28 30 29 1 2 3 5 4 6 7 8
	SO	1 2 3 5 4 6 7 8 18 19 20 21 22 23 24 25 26 27 28 29 30
Pop5	NSO	5 4 3 2 1 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
	SO	1 2 3 4 5 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30

We can see that for this complicated data set, separate analysis of each population, or non-synchronized ordering, resulted in wrong map (NS-orders). By contrast, consensus analysis via synchronized ordering resulted in correct maps (S-orders).

6.2 QTL Mapping

A huge amount of mapping results for cereal crops is spread in the literature, and for some species (e.g., barley and wheat) considerable efforts have already been invested to summarize this information. Low accuracy of estimated chromosomal positions of detected QTLs is one of the major obstacles in application of QTL mapping results for marker-assisted selection and positional cloning of genes corresponding to the QTLs. Here we discuss the advantages of *multiple* approach of QTL analysis that includes multiple-trait or multiple-environment models combined with multiple-interval mapping. These combinations of methods proved very efficient, especially for increasing QTL detection power and mapping precision. One of the new interesting applications is “expression-

QTL” mapping using microarray data. A wide spectrum of questions about the nature of quantitative variation considered in terms of “genetic architecture” can be addressed by *multiple* approaches, including genomic distribution of QTL effects; contribution of additive and non-additive effects to genetic variation and response to selection; role of overdominance, epistasis, coadaptation, and pleiotropy in heterosis and speciation; developmental variation in QTL effects, and genetic basis of “reaction norm” (QTL-E interactions). Clearly, the complicated nature of quantitative variation requires adequate, and usually rather complicated, analytical tools despite a researcher’s natural preference to employ simple tools. The cost of generating experimental data by far exceeds the cost of data analysis (Churchill and Doegre 1994). To our opinion, any complication of mapping models is justified if it increases QTL detection power and mapping accuracy. Consideration of the wide spectrum of questions and analytical challenges related to modern QTL mapping would require a separate volume; hence we consider here shortly only two aspects: multiple-trait and multiple-environment analyses.

6.2.1 Multiple Trait Analysis

The detection power and mapping resolution of marker analysis of quantitative trait variation is a key factor affecting various applications of mapping quantitative trait loci (QTL). These characteristics strongly depend on the effect of the QTL in question relative to the phenotypic variance of the trait in the mapping population. As shown by Lander and Botstein (1989), the expected value of the log-likelihood test statistics increases monotonically with the contribution of the QTL to the broad sense heritability (H^2):

$$\text{ELOD} = -1/2 N \log(1 - H^2), \quad (6.1)$$

where N is the sample size. Several strategies have been proposed to improve the precision of QTL mapping, based on new experimental designs to suit specific mapping goals and an organism’s breeding system, and new QTL mapping models and algorithms to extract maximum information about QTL locations and effects. A spectrum of proposed mapping designs include replicated progeny testing, selective genotyping and selective DNA pooling, advanced intercross line, and recombinant inbred line analysis, congenital lines analysis for non-selfing organisms (e.g., mouse), etc. Among various analytical schemes of special interest is multiple interval mapping (MIM) (Kao et al. 1999) that allows improving QTL mapping by reduction of the residual variation, i.e., taking into account the effects of co-segregating QTLs. A complementary situation, when one QTL (or a chromosome segment) affects simultaneously several traits, can also be considered, resulting in increased resolution power (Korol et al. 1987, 1995, 2001; Jiang and Zeng 1995; Ronin et al. 1995, 1999). Indeed, in QTL

mapping, the experimental design usually includes simultaneous measurements of many related and unrelated quantitative traits and subsequent treatment of the individual traits.

In the simplest case of two non-correlated traits, the advantage of joint analysis is in the increase of the “*multivariate effect*” according to $d^2 = (d_x/\sigma_x)^2 + (d_y/\sigma_y)^2$, where d_x and d_y are the substitution effects of the QTL for traits x and y , and σ_x and σ_y are the corresponding standard deviations within the QTL groups (i.e., residual standard deviations). Consequently, for a population with 1:1 ratio of the alternative QTL groups (like backcross, dihaploid, or RIL) the bivariate analogue of H^2 can be represented in the form

$$H_{xy}^2 = 1/4d^2/(1 + 1/4d^2) \quad (6.2)$$

The situation becomes more complicated when correlated traits are involved. It can be shown (Korol et al. 1995) that Eq. 6.1 remains valid in analysis of correlated traits:

$$\text{ELOD}(x, y) = -1/2 N \log(1 - H_{xy}^2) \quad (6.2a)$$

with $H_{xy}^2 = 1 - T_1/T_2$, where $T_1 = \sigma_x^2 \sigma_y^2 (1 - R_{xy}^2)$, and $T_2 = (\sigma_x^2 + 1/4d_x^2)(\sigma_y^2 + 1/4d_y^2) - \sigma_x^2 \sigma_y^2 [R_{xy} + d_x d_y / (4\sigma_x \sigma_y)]^2$. It was shown earlier that either $\text{ELOD}(x, y) \geq \text{ELOD}(x)$ and $\text{ELOD}(x, y) \geq \text{ELOD}(y)$ follows from $H_{xy}^2 \geq H_x^2$ and $H_{xy}^2 \geq H_y^2$, respectively (Korol et al. 1995; Ronin et al. 1999). Therefore, the increment in H_{xy}^2 , compared with H_x^2 , will result in an increased resolution of the mapping analysis, no matter how this increment in H_{xy}^2 was produced: due to (i) the pleiotropic effect of the QTL on x and y ; (ii) residual correlation between x and y (within the QTL groups) caused by non-genetic effects or segregation of unlinked QTLs; or (iii) the *combined effect* of both factors (i) and (ii). In other words, instead of separate analyses of traits x and y , one can conduct joint analysis of x and y , that is formally equivalent to transformation of two-dimensional phenotype into one-dimensional phenotype. For the new phenotype, a higher ratio of the between QTL-group difference to the residual variation can be achieved due to pleiotropic QTL effect on both traits, and residual correlation between the traits caused by non-genetic factors and segregation at other QTLs.

The increased number of parameters to be estimated complicates the application of this approach when a large number of traits are considered. With n traits and single interval mapping, the model should include $(n^2 + 5n + 2)/2$ parameters (QTL position, n mean values, n effects, n residual variances, and $n(n-1)/2$ covariances); which amounts to 76 parameters for $n=10$ in the simplest case of RIL or dihaploid population. One possibility is space transformation by using the principal component analysis (PCA) (Weller et al. 1996; Mangin et al. 1998) applied to the trait complex across the entire data set. Although this approach seems very attractive, it cannot directly solve the

problem when the population segregates for more than one QTL, especially if some effects are relatively strong (see discussion in Korol et al. 2001).

Our approach is also based on transformation of the initial trait space into a space of a lower dimension. In the simplest case of a single QTL analysis of a RIL (dihaploid) mapping population, the resulting space is one-dimensional independent of the number of traits, whereas two-QTL analysis for such a population will employ a two-dimensional model (for F_2 these will be 3- and 8-dimensional models, correspondingly). The main difference of this approach from the foregoing PCA-based models is in the fact that the residual variance-covariance matrix is considered interval-dependent: its elements are a subset of the vector of unknown parameters to be estimated by the employed procedure *for each interval*, so that for QTLs residing in different genomic segments the resulting (transformed) traits could be very different. This interval-dependence remains a notable characteristic of our multivariate algorithms implemented in MultiQTL software.

Application of multiple-trait (MT) analysis is illustrated using data on a *T. durum* \times *T. dicoccoides* F_2 mapping population with some morphological traits related to domestication (Peng et al. 2003). Out of 11 scored traits we employed here a subset of traits related to plant growth and productivity: plant height, heading date, grain yield, spike weight per plant, single spike weight, 100-grain weight. Each genotype was characterized by the trait scored as average over its selfed progeny F_3 . Significant effects were found for this multi-trait complex on chromosomes 1B, 2A, 4A, 5A, and 6B. The example in Table 6.2 demonstrates some advantages of MT analysis and its combination with MIM. MT mapping appears to be more accurate compared to single-trait

Table 6.2 Comparing single interval and multiple interval mapping in multiple trait analysis (data on F_2 *T. durum* \times *T. dicoccoides*, Peng et al. 2003)

Chr	Q	SIM					MIM			
		LOD	m_L	m_L^*	ΔLOD	$p(e = 0)$	LOD	m_L	ΔLOD	$p(e = 0)$
1B	Q1	21.4	16.6	20.0	9.9	0.11	42.4	3.7	24.1	<0.001
	Q2		26.5	26.9				5.6		
2A	Q	13.2	14.7	17.5			25.5	2.7		
4A	Q1	16.2	9.0	19.8	6.4	0.56	30.1	6.2	18.8	<0.001
	Q2		20.9	64.0				11.2		
5A	Q1	12.7	22.2	50.2	3.3	0.89	27.7	9.3	12.2	0.019
	Q2		61.9	98.1				15.0		
6B	Q	12.9	16.5	24.9			24.3	2.3		

Notes: SIM and MIM – simple and multiple interval mapping; *LOD* – lod value for linked-QTL model (for 1B, 4A, and 5A) and single-QTL model; m_L and m_L^* – standard variation of QTL position using, respectively, MT analysis and single-trait analysis for the trait with highest pleiotropic effect of the considered chromosome; ΔLOD – lod decrement upon replacement of model with epistatic interaction between linked QTL ($e \neq 0$) by $e = 0$ model; $p(e = 0)$ – significance of epistasis in the chromosome for the considered multiple-trait complex.

analysis even if the comparison is with trait with highest effect of the chromosome in question. Further increase in efficiency of mapping analysis is by reducing the residual multitrait variation by moving from MT-SIM to MT-MIM. This combination provides a several-fold improvement of mapping accuracy. Another important advantage is increased discrimination between alternative hypotheses about trait genetic architecture. Of special interest is testing for the presence of epistasis. As can be seen from the presented example, in addition to better mapping accuracy, MT-MIM analysis may considerably improve detection of epistasis.

Clearly, in conducting these types of analysis one would like to know what are the individual effects of the detected and mapped QTL on each of the traits, are all of them significant, should all the traits be maintained in the trait complex or some of the traits are irrelevant for the considered chromosome, i.e., do not contribute to QTL detection power and mapping accuracy. The first of these problems is easily solved by “projection” of the multivariate effect on the axes of initial traits (Korol et al. 2001). To address the last question we employ a modified permutation test: for each traits reshuffling of the trait values are conducted relative to the set of the remainder traits and the marker set. This allows testing (a) the trait “contribution” to the LOD, and (b) the significance of the putative QTL additive and/or heterozygous effects.

6.2.2 Paradoxical Consequences of Variance-Covariance Effect

In traditional QTL analysis, one standard hypothesis is homogeneity of residual variance in the QTL groups. This simplifying assumption not always fit real situations: the alternative QTL groups may differ with respect to their residual variances (heteroscedasticity); such QTL could be referred to as variance effect QTL (Korol et al. 1996). Thus, allele substitution at a QTL may cause: (a) change in trait mean value (standard situation), (b) change in residual variance, and (c) both effects. The question of how the resolution power of multiple-trait analysis depends on the effect of the QTL on the residual variation of the traits. It concerns the effects on residual variance and residual correlation. The problem of identification QTL effects on trait variances can be discussed for both single- and multiple-trait analysis (Ronin et al. 1995; Korol et al. 1995). Its importance extends beyond pure analytical aspects. Indeed, the effect of a QTL on the variance is sometimes economically more critical than on the mean (e.g., for earliness, flowering time, ripening time under machine harvesting, time to hatching in chicken). The same applies to QTLs related to fitness traits in natural populations, e.g. seed dormancy, or flowering time. In general, variance effect QTLs may be significant players in the genetic control of plasticity, including plasticity of such quantitative traits as gene expression levels scored in microarray assays (see examples below).

Consider a dihaploid population. For a putative QTL (say, A/a) one can compare two situations with respect to the residual variance in the QTL groups:

(i) $\sigma^2_{AA} = \sigma^2_{aa} = \sigma^2_0$ and (ii) $\sigma^2_{AA} \neq \sigma^2_{aa}$. The assumption of equal variances (homoscedasticity) is a usual simplification in ANOVA and QTL mapping. It was found earlier, that if the variance effect of A/a is large enough, then, despite of the increased number of parameters, the more correctly specified model (that allows for variance effect) provides an increase in the resolution power, as compared to the situation (i). This is not unexpected, if either σ^2_{AA} or σ^2_{aa} in (ii) is lower than σ^2_0 from (i). But this effect holds even if $\sigma^2_{AA} > \sigma^2_{aa} = \sigma^2_0$ or $\sigma^2_{aa} > \sigma^2_{AA} = \sigma^2_0$, which on the first glance is much less expected. Earlier we showed (Korol et al. 1996) that expected LOD value can be presented as a function of scaled QTL substitution effect ($\delta = d/\sigma_{aa}$), the ratio of the residual variances (α^2), and sample size N , in the form:

$$\text{ELOD} = L(\alpha, \delta) = N \log[0.5(\alpha + \alpha^{-1}) + 0.25\delta^2/\alpha]. \quad (6.3)$$

Expression (6.3) shows a monotonic increase with distance between the trait means of the QTL groups (δ) and a non-monotonic (in general) dependence on the variance ratio of the groups (α). A similar question about the dependence of detection power and accuracy of parameter estimates on the assumption of equal variance-covariance matrices was considered in the multi-trait analysis (Ronin et al. 1995). For simplicity, one may assume that no such effects are presented in the data and put $\Sigma_{aa} = \Sigma_{AA}$, where symbol Σ denotes the residual variance-covariance matrices. However, in both single- and multi-trait analysis, such simplifications may lead to a loss in the detection power and precision of the estimates, if these assumptions do not fit the data. On the other hand, if indeed $\Sigma_{aa} \neq \Sigma_{AA}$, then the resolution could be improved significantly if this fact is taken into account (Ronin et al. 1995).

The foregoing consideration is illustrated using data on QTL analysis of gene expression in *Arabidopsis thaliana* (West et al. 2007). By developing microarray chips on mapping populations it is now possible to map QTL involved in regulation of gene expression. Thousands of eQTL have been mapped in various crosses in mice, rat, arabidopsis, maize, human, yeast. To test for variance effect we employed the publicly available data from the cited study (data repository at <http://www.ebi.ac.uk/arrayexpress/under> nos. E-TABM-126 and E-TABM-224). Analysis of a small sub-sample of 200 genes from this dataset revealed various patterns of eQTL effects, including strictly “traditional” QTL affecting only mean expression level of a gene (majority of eQTL), strictly variance-effect QTL (minority of eQTL), and QTL displaying both effects (see Lacaze et al. 2009a). Some of eQTL of the third category appeared to be highly pleiotropic, affecting many genes. An example of such eQTL from chromosome 2 is presented in Table 6.3. The example in Fig. 6.4 illustrates the same idea for trait pairs. For the chosen pair (genes 257054_at and 267002_s_t) the eQTL on chromosome 2 affects simultaneously the expression level of both genes, residual variance of 257054, and residual correlation between the traits (LOD = 26.34). Remarkably, this eQTL affects similarly

Table 6.3 Detecting variance effect eQTL in *Arabidopsis thaliana*

Trait (gene)	eQTL position		LOD	Testing the effect					
				m		V		mV	
	Chr	L, cM		p	m _L	p	m _L	p	m _L
246481-s-at	3	1.8	9.7	<0.001	3.6	0.54	18.9		
259361-at	1	52.9	4.2	<0.001	7.4	0.42	24.2		
262609-at	5	41.2	5.5	<0.001	12.7	0.88	41.2		
250995_at*	2	36.3	5.7	0.35	15.4	<0.001	5.0		
246068_at	4	48.8	3.2	0.72	21.7	<0.004	11.1		
257054_at*	2	37.4	8.0	0.001	9.5	<0.001	2.0	<0.001	1.4
259625_at	2	13.7	6.6	0.006	12.9	<0.001	14.2	<0.001	10.5
247073_at*	2	35.2	10.7	<0.001	2.2	<0.001	4.4	<0.001	3.6

Notes: LOD – lod value for the general model allowing for variance effect; *p* and *m_L* – significance of eQTL and standard variation of the eQTL mapped position for the corresponding model (*m* – no variance effect, *V* – only variance effect, and *mV* – general model). The best model for the corresponding traits is presented in bold font. Note that the three examples marked by asterisk represent a pleiotropic eQTL that displayed a variance effect (with or without simultaneous effect on the mean expression level) for many other traits as well.

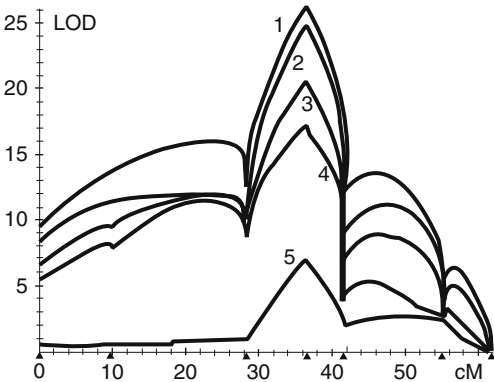


Fig. 6.4 Variance-covariance effect in two-trait eQTL analysis on chromosome 2 of *A. thaliana* RIL population (expression of genes 257054_at and 267002_s_t, respectively): **(1)** general model (allowing for effect on mean value and variance effect for both traits, and covariance effect); **(2)** and **(3)** no variance effect for trait 1 and 2, respectively; **(4)** no covariance effect; **(5)** two-trait model with (2), (3), and (4). The effects on mean value of the two traits were significant at $p=0.002$ and $p<5\cdot10^{-4}$, respectively. Comparing the models gives: $p\{(2) \text{ vs } (1)\}>0.25$; $p\{(3) \text{ vs } (1)\}<5\cdot10^{-4}$; $p\{(4) \text{ vs } (1)\}<<5\cdot10^{-5}$; $p\{(5) \text{ vs } (1)\}<<10^{-5}$ [using data of West et al. (2007)]

many other genes and gene combinations. This means that genetic plasticity of gene expression (defined as variation of expression across genetic backgrounds), as well as coordination of expression of different genes across genetic backgrounds, can be under control of the same eQTL (Lacaze et al. 2009a). The examples in Fig. 6.4 and Table 6.3 illustrate also a technical advantage of testing

variance effect: including it into the mapping model (when proved significant) improves the significance and the accuracy of the estimated eQTL position, as well as the power of eQTL detection (not shown).

6.2.3 *Multiple Environments*

Differential expression of a phenotypic trait by genotypes across environments, or genotype \times environment ($G \times E$) interaction, is an old problem of primary importance for quantitative genetics and its applications in breeding, conservation biology, theory of evolution. Successful attempts to genetically dissect quantitative variation employing molecular markers have shifted the focus of $G \times E$ interaction analysis from the genotype to gene level in various QTL \times E models (Hayes et al. 1993; Jansen et al. 1995; Romagosa et al. 1996; Sari-Gorla et al. 1997). For breeding purposes the primary concern is possible environmental instability in the effect of mapped QTLs that might become candidates for marker-assisted selection. To evaluate stability of QTL manifestation in crop species quite a few immortal mapping populations have been developed for trait scoring under various environmental conditions (Hayes et al. 1993). In addition to testing the hypothesis of QTL \times E interaction, simultaneous treatment of data from multiple environments provides a significant increase in statistical power of QTL detection and accuracy of the estimated QTL position and effect (Jansen et al. 1995). However, such an analysis is limited by situations where the environments can be obviously characterized by some parameters, like day length, or irrigation-fertilization treatments, etc. (like ‘fixed effects’ model of ANOVA). When these characteristics are not available the application of the “general” QTL \times E mapping model (see below) is accompanied by increased number of parameters involved in the model. In such a case one could think about a ‘random effects’ model, so that the number of parameters for each QTL will include only main effect, the variance of QTL \times E interaction and QTL position.

Although the last option seems to be very attractive, it has its own drawbacks especially if we are going to deal with environmental variation associated with different localities. Indeed, some (or many) localities may manifest quite repeatable differences from each other, justifying the ‘fixed model’ approach. Moreover, the information of geographically-specific QTL effects may be of practical importance. In such a case, a fixed-effect model, whenever possible, is preferable over the random-effect model because the latter hides the biological (geographic) specificity of the QTL effect, compressing all the results to an estimate of variance.

Earlier, we developed an approach of QTL mapping analysis allowing for QTL \times E interaction across a large (in fact, unlimited) number of environments, without the necessity for a corresponding increase in the number of parameters (Korol et al. 1998). The proposed approach is especially relevant in situations of

geographic variation of external conditions, where the fixed model approach is desirable but not always easy to implement. Its main distinction is in the chosen way of approximation of the dependence of putative QTL effects on environmental states. We hypothesized that environmental dependence of a putative QTL effect can be represented as a function of environmental mean value of the trait. In reality, each environment is a complex of abiotic (temperatures, humidity, ion concentrations, etc.), biotic (parasites, pathogens, competitors, etc.) and agro-technological features. These features could strongly affect the manifestation of quantitative traits and the effects of QTL, but are difficult to characterize quantitatively. As first suggested by Eberhard and Russel (1966), we advocate that the measured trait values of the mapping population (e.g., trait means) may serve as objective integral characteristic of the environmental state. Accordingly, a larger number of traits should provide a better “bioindication”. In the simplest form, one can approximate the environmental dependence of the effect of allele substitution at a QTL by a polynomial over the mean values of the same trait across the environments (Korol et al. 1998), which can be considered as an extension of the joint regression model proposed by Finlay and Wilkinson (1963).

Consider a simplified situation when the trait of interest (x) of dihaploid populations depends on a single QTL, Q/q . For an arbitrary genotype j the trait measurement in the i th environment can be presented as

$$x_{ij} = m_i + 0.5g_k a_i + e_{ij}, \quad (6.4)$$

where m_i is the mean trait value in the i th environment, g_k is either $+1$ (for QQ genotypes, $k=1$) or -1 (for qq genotypes, $k=2$), a_i is the effect of allele substitution at putative QTL on trait in environment i , and e_{ij} is a random variable with zero mean and variance σ^2_i . If we find $a_i \sim a$ for any i , then no $G \times E$ interaction is manifested by Q/q . Assume that Q/q resides in some interval $(k, k+1)$ of a chromosome marked by a series of marker loci, M_j/m_j , with recombination rates r_1 and r_2 in $M_k/m_k - Q/q$ and $Q/q - M_{k+1}/m_{k+1}$, respectively. Then, the expected distribution of the trait x in each of the four marker groups $U_{mkmk+1}(x) = U_1(x)$, $U_{Mkmk+1}(x) = U_2(x)$, $U_{mkMk+1}(x) = U_3(x)$ and $U_{MkMk+1}(x) = U_4(x)$ can be written as mixtures:

$$U_i(x) = \pi_i f_{qq}(x) + (1 - \pi_i) f_{QQ}(x), \quad i = 1, \dots, 4 \quad (6.5)$$

with mixture proportions $\pi_i = \pi_i(r_1, r_2)$ dependent on r_1 and r_2 . Here $f_{qq}(x)$ and $f_{QQ}(x)$ are the trait distributions in the QTL groups qq and QQ , respectively.

According to the proposed approach, the unknown effects a_i and the residual variances σ^2_i are represented by low degree polynomials (usually, 2 or 3). Clearly, the main question remains: to what extent the “bioindicating” trait, or a (linear) combination of traits, will indeed be informative with respect to the dependence of the target QTL effect on environmental states. No prior answer is possible, but analysis of real data on barley demonstrated (Korol et al. 1998)

that such an assumption is quite realistic even with the simplest univariate mode of the approximation, $a_i = f(m_i)$. In other words, in the log-likelihood analysis based on model (4), we could replace the corresponding coordinates of the genetic parameters by polynomials:

$$\begin{aligned} a_i &= \alpha_0 + \alpha_1 m_i + \alpha_2 m_i^2 + \dots + \alpha_s m_i^s, \\ \sigma_i^2 &= \beta_0 + \beta_1 m_i + \beta_2 m_i^2 + \dots + \beta_t m_i^t. \end{aligned} \quad (6.6)$$

Thus, instead of estimates of r_1 , a_i and σ_i^2 , the ML-procedure will provide estimates of r_1 and regression coefficients $\alpha_0, \alpha_1, \dots, \alpha_s$; β_0, β_1 , and β_t . The final degrees s and t are to be chosen on the basis of maximum statistical significance of the hypothesis ‘QTL \times E interaction’ versus the alternative of ‘no QTL \times E interaction’.

The approximation (6) can also be represented in form of deviations from the mean values of the trait m averaged over environments, i.e. with terms $\alpha_k(m-m_i)^k$ and $\beta_j(m-m_i)^j$ instead of $\alpha_k m_i^k$ and $\beta_j m_i^j$. Then, the genetic interpretation of the coefficient α_0 is that it specifies the average substitution effect at the putative QTL, whereas coefficients α_i ($i > 0$) reflect the stability of the QTL effect over environments. This parallels the stability analysis of Eberhard and Russel (1966), although they used only linear regression in their model. Different reasons could be suggested why linear approximation for QTL dependence on environment may be not sufficient. For example, one may mention the effect of canalization of gene effects (Korol et al. 1994). Likewise, the coefficient β_0 specifies the residual variance under average conditions, whereas β_i ($i > 0$) reflect the stability of the residual variance with deviation from average conditions. These considerations are of interest in the genetic dissection of phenotypic plasticity (Lacaze et al. 2009b). Another approach to analyze QTL \times E interaction without direct specification of the ‘physical’ characteristics of the environments was proposed by Romagosa and co-authors (1996). Their algorithm is based on clustering the environments using a few detected QTLs with most variable effects across environments. In fact, this is another form of the same general idea of *bioindicators* as a tool for characterizing “anonymous” environments.

Although the application of the above bioindicator regression model was already illustrated on real data (see Korol et al. 1998), here we will provide some examples of QTL \times E analysis, based on a more standard model of Jansen et al. (1995) implemented in MultiQTL package. Assuming that locus Q/q belongs to interval $(k, k+1)$, log-likelihood for a sample of measurements x_{ijs} in marker groups with sizes N_s ($s = 1, \dots, 4$) can be written as:

$$\ln(\vartheta_1) = \sum_{sij} \ln U_s(\vartheta_1, x_{ijs}) = \sum_{sij} [\pi_s f_Q(x_{ijs}) + (1 - \pi_s) f_q(x_{ijs})], \quad (6.7)$$

where vector $\vartheta = \vartheta_1 = (r_1, m, a, \sigma^2)$ of unknown parameters specifies the putative QTL position in the considered interval, population means (m_1, m_2, \dots, m_n), QTL effects (a_1, a_2, \dots, a_n), and residual variances ($\sigma_1^2, \sigma_2^2, \dots$,

σ_n^2) in each environment. Clearly, vector $\boldsymbol{\vartheta}_1$ defines the hypothesis H_{1E} that there is a fluctuating QTL effect across environments (QTL \times E interaction) associated with the considered marker interval. For each pair of flanking markers, maximization of the likelihood function is conducted using sliding scanning across “trial” positions r_{trial} of the putative QTL between the flanks, resulting in estimates of parameters μ_i , a_i , and σ_i^2 for each r_{trial} . The assumption of a significant QTL effect with no QTL \times E interaction (H_{10} hypothesis) can be parameterized by a vector $\boldsymbol{\vartheta} = \boldsymbol{\vartheta}_{10} = (r_1, m, a^*, \sigma^2)$, where $a = a^*$ represents the condition $a_1 = a_2 = \dots = a_n$. The existence of a QTL (H_1 versus H_0 or H_{10} versus H_0) is tested by permutation tests (Churchill and Doegre 1994).

In order to detect a QTL showing QTL \times E interaction, the hypothesis of constant substitution effect of the QTL across environments should be compared to the more general hypothesis of varying effect across environments. For that, interval QTL analysis allowing for QTL \times E interaction (H_{1E} model) is performed and followed by fitting also a model assuming no QTL \times E interaction (H_{10} model). We then simulate multiple data sets (e.g., 1000) using as the basis the parameter values estimated from the H_{10} model. For each of these sets, a QTL analysis is performed with and without QTL \times E assumption. The difference in LOD score of the two models is employed as a criterion of the significance of QTL \times E interaction. Namely, we count the proportion of cases where the difference in LOD score between the two models on simulated data is higher than the difference obtained on real data.

For illustrations we employed data on barley Steptoe-Morex dihaploid population scored across multiple environments (Hayes et al. 1993). In addition to the foregoing basic analysis, we show also some more complicated examples, including two-trait analysis across multiple environments and QTL \times E combined with MIM analysis. Table 6.4 demonstrates several of the aforementioned aspects. It allows comparing the results of joint analysis for the trait scores across nine environments with those for the environment with the strongest detected QTL effect for the considered chromosome. It can be seen that with both SIM and MIM, ME-analysis provides a considerable improvement of mapping accuracy, one of the major challenges in QTL mapping. Another important point is the contribution of MIM analysis in detecting QTL \times E interaction. Four out of five detected QTL displayed such interaction with MIM analysis, and only two with SIM.

Next example, on grain protein, is about epistasis in multiple-environment analysis. Two QTL are detected in chromosome 2H for this trait. The hypothesis of no QTL \times E interaction is rejected at high level of significance for both Q1 and Q2 either with SIM and MIM analysis. Significant ($p = 0.003$) epistasis is detected between Q1 and Q2 with MIM (but not SIM) analysis, whereas epistasis \times E interaction proved non-significant. Thus, the environment-specific estimates of epistasis $e_i \pm m_{ei}$

$$\begin{array}{cccccccccc} e_i & -0.025 & -0.213 & -0.197 & -0.211 & -0.137 & -0.108 & -0.249 & -0.130 & -0.383 \\ m_{ei} & 0.175 & 0.119 & 0.080 & 0.082 & 0.076 & 0.081 & 0.109 & 0.124 & 0.098 \end{array}$$

can be replaced by overall estimate $e \pm m_e = -0.194 \pm 0.036$.

Table 6.4 Comparing single- and multiple-environment analysis for barley data on plant height scored in nine environments (Hayes et al. 1993)

Chr	<i>L</i> cM	Single-environment (SE)				Multiple-environment (ME)						
		SIM		MIM		SIM		MIM				
		<i>m_L</i>	<i>p</i>	<i>m_L</i>	<i>p</i>	<i>m_L</i>	ΔLOD	<i>p</i> (<i>a</i> × <i>E</i>)	<i>m_L</i>	ΔLOD	<i>p</i> (<i>a</i> × <i>E</i>)	
2	Q1	41	3.6	<10 ^{−5}	2.5	<<10 ^{−7}	2.7	8.9	<10 ^{−4}	1.8	13.3	<10 ^{−5}
	Q2	165	14.4	0.002	13.2	<0.001	5.5	0.7	0.60	7.0	1.0	0.73
3		54	2.9	<10 ^{−5}	1.4	<<10 ^{−7}	1.6	8.0	<10 ^{−4}	1.4	15.9	<<10 ^{−5}
5		73	11.0	<0.001	7.9	<10 ^{−5}	1.7	1.7	0.13	1.2	4.2	0.002
7		93	12.8	<0.001	10.8	<10 ^{−5}	8.6	1.1	0.20	4.5	3.4	0.012

Notes: SIM and MIM – simple and multiple interval mapping; *L* – QTL position; *p* and *m_L* – significance of QTL effect and standard variation of *L* for corresponding model (*p* values are not shown for the ME models to save place: in all of presented cases they were $<<10^{-6}$). Δ LOD – lod decrement upon replacement of QTL \times E model by QTL = const model; $p(a \times E)$ – significance of the QTL \times E interaction.

The last example is about analysis of pleiotropic QTL in the multiple-environment analysis. We consider simultaneously two traits, grain protein and malt extract, that seem to be affected by the same regions on chromosomes 1H and 2H: one QTL on 1H (the corresponding model is referred to as 1Q) and two QTL on 2H, one for both traits and the other one only for gp (model 2Q). The efficiency of *multiple* approach compared to the standard scheme is displayed in a considerable increase in QTL detection power and improvement of mapping accuracy (decrease in m_L) (Table 6.5). The increased mapping precision can be seen as a general tendency: within standard SE + SIM analysis, joint analysis of two traits is superior to single trait analysis; within single trait analysis, ME is superior to SE, and ME combined with MIM is better than ME + SIM combination. In fact, due to *multiple* approach we could reach $m_L \sim 1\text{--}3$ cM, i.e., about an order of magnitude lower compared to standard analysis with the data from the environment with the strongest QTL effect among the nine considered. However, although increase in accuracy is a general tendency of multiple solutions, some exclusions may also be found (compare the cases Me + SIM and Me + MIM for Q2 on chromosome 2 in Table 6.4). It is hard to predict such cases before the analysis; this means that a lot of work still remains to be done to develop the theory properly.

To illustrate the last statement we would like to return back to the question of multiple marker analysis in QTL mapping. In order to improve the accuracy of interval mapping, a procedure called “composite interval mapping” (CIM) was developed (Zeng 1994; Jansen and Stam 1994). Its idea is similar to MIM analysis: if the analyzed chromosome carries more than one QTL, the accuracy of mapping of each QTL can be increased if the information provided by its flanking markers is complemented by additional (carefully selected) markers of the same chromosome. In such a case, CIM can be considered as a powerful tool for multiple-QTL analysis. However, if the chromosome carries only one QTL, employment of CIM will, probably, not improve the situation. Still, the idea of taking advantage of information provided by all scored markers seems promising, especially for situations of “information deficit”, e.g., upon very limited sample size, missing marker scores, score errors, etc. One of the ways to implement this idea would be fitting a multiple regression model based on the fact that the expected effect of marker m is proportional to $(1 - 2r_m)$, where r_m is the recombination rate between the marker m and the QTL (for more details of such approach see: Kearsey 1998, Ronin et al. 1999, and the next section). Such whole-chromosome mapping (WCM) can also be extended to situations with two or more linked QTLs.

The following simulated example with single-QTL data shows how WCM approach can improve the mapping accuracy compared to standard interval mapping. We compared WCM, SIM, and CIM in two situations: (a) with fully genotyped data, (b) under 8% of missing marker scores. The analysis was conducted using three packages: a home made software (MultiPool) for WCM, MultiQTL for SIM, and QTL Carthographer for CIM. Single backcross population of 250 individuals with 14 markers on one chromosome of 120 cM was simulated by MultiQTL assuming Haldane mapping function. Trait values of individuals were simulated as normally distributed random variable with

Table 6.5 Joint analysis of two traits, grain protein (gp) and malt extract (me) across multiple environments

Model	SE + SIM				ME + SIM				ME + MIM			
	gp or me		gp & me		gp or me		gp & me		gp or me		gp & me	
	m_{L1}	m_{L2}	P	m_{L1}	m_{L2}	P	m_{L1}	m_{L2}	P	m_{L1}	m_{L2}	P
gp 1Q	21		100			3		100		3		100
gp 2Q	21	33	100	3		100	3	100	2	1	100	100
me 1Q	15		60	22	40	100	18	69	4	3	88	100
me 2Q	23	44	65			9	38	100	3	34	100	

Notes: The QTL mapping accuracy (standard deviation of the estimated position m_L (cM)) and QTL detection power (P) at significance level 0.01 are shown. The results of single-trait (gp or me) and two-trait (gp & me) analyses are in columns 2,4,6 and 3,5,7, respectively.

variance 1 and mean depending on genotype at single QTL with additive effect $d=0.75$ situated in the middle of the eighth marker interval. To compare the methods for accuracy, the size of the confidence interval of mapped QTL position is employed. To compare SIM and WCM, we calculated standard deviations of QTL position (SD_L) based on bootstrap resampling (no such option was found in current QTL Carthographer software), whereas for SIM vs CIM comparison, the size of the confidence interval based on 2 LOD drop (CI_{2LOD}) was employed (no such option is possible in WCM).

In case (a), genotypes of markers outside the interval containing QTL give no additional information about QTL effect and position. Hence, SIM gives accurate estimation. For this situation, SD_L values for SIM and WCM (based on bootstrap resampling) were about 6.3, whereas CI_{2LOD} for SIM and CIM were about 24 cM. In case (b), neighbor markers outside of the interval containing QTL give some additional information about QTL position that comes from individuals where one or both flanking markers are missing. Nevertheless, changes in the haplotype structure of the mapping population (due to data missing) caused some changes in linkage relationships between markers and QTL. This leads to appearance of false peaks of LOD function fitted in SIM and CIM (see Fig. 6.5), which is not the case for fitted WCM model. Hence, WCM become more effective than SIM and CIM. Namely, in the simulated example, SD_L value for WCM (6.5) was less than that for SIM (9.2), whereas CI_{2LOD} for SIM and for CIM both were about 33 cM.

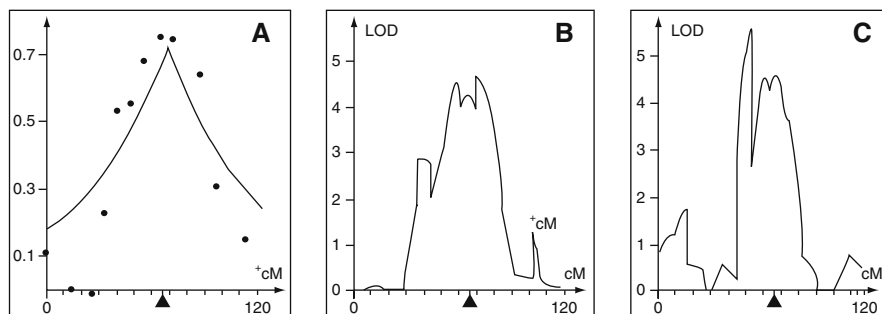


Fig. 6.5 Three approaches of QTL analysis (WCM, SIM and CIM): a simulated example with single QTL in dihaploid population with missing genotypes. Position of the simulated QTL is marked by black triangle. **(A)** Observed single markers effects of on trait (bold dots) and expected (continuous curve) effect for each chromosomal position in situation when QTL effect and position were determined using WCM. Positions of markers and chromosome length were estimated using the observed recombination rates. Although deviations of the observed marker effects from the expected ones are rather high (because of missing data), the accuracy of the estimated QTL position estimation is good. **(B)** LOD graph in SIM. Local false peaks appear as a result of population structure disturbance induced by missing data. **(C)** LOD graph in CIM. False peaks are even higher than in SIM because the effect in every single interval of chromosome is calculated by reducing the residual variation due to co-factor "effects" of a selected subset of linked markers

6.3 High-Resolution Mapping Based on Selective DNA Pooling

One of the major conditions for efficient association analysis is large samples size. If applied to genome-wide scanning based on SNP chips, this implies very high project cost (in human genetics, some data sets for association analysis included $\sim 10^3$ – 10^4 individuals). Selective DNA pooling (SDP) analysis may be a solution to this problem (Darvasi and Soller 1994). SDP is a cost-effective means for a scan for linkage between markers and QTLs in suitable populations. It is based on scoring marker allele frequencies in DNA pools from the tails of the population trait distribution (Hillel et al. 1990). Various analytical approaches have been proposed for QTL detection using data on multiple families with SDP analysis (Dekkers 2000; Carleos et al. 2003; Brohede et al. 2005; Johnson 2005; Wang et al. 2007). Of special interest will be a new experimental procedure, Fractioned Pool Design (FPD), aimed to increase the reliability of SDP mapping results, by “fractioning” the tails of the population distribution into independent sub-pools (Korol et al. 2007b). It may be advantageous for pooling analysis with SNPs. Combining SNP microarray analysis with DNA pooling can dramatically reduce the cost of screening large numbers of SNPs on large samples, making chip technology readily applicable for genome-wide association mapping. We discuss here the prospects of FPD extension to LD-mapping using natural populations of wild progenitors of *Triticeae* cultivated species.

6.3.1 Standard Selective DNA Pooling Approach to QTL Mapping

The experimental material for QTL mapping based on SDP consists of individuals selected from the tails of the mapping population trait distribution. We designate as positive and as negative QTL alleles the alleles increasing and decreasing trait value respectively. Then the frequency of the positive QTL allele will be higher in the group of individuals having high trait value and lower in the group of individuals having low trait value; the opposite will be true for the negative QTL allele. Due to linkage, this difference in the frequency of the positive and negative QTL alleles in groups with high and low trait values produces a parallel difference in the frequency of marker alleles that are in coupling linkage phase to these QTL. Analyzing population marker-allele frequency differences at several marker loci enables estimating the chromosomal position of the QTL. In fact, the analysis can be conducted on a single mapping population (e.g., RIL family) or simultaneously on a set of populations.

It is convenient to denote the pools as high (H) and low (L), respectively, and the two alternative SNP alleles at the linked marker locus m ($m = 1, \dots, M$) as alleles A_m and B_m , respectively. Using this notation, we define the statistic D_m as a characteristic of allele divergence in the two tails,

$$D_m = |FHA_m - FLA_m|, \quad (6.8)$$

where FHA_m and FLA_m are the frequencies of allele A_m in the high and low pools, respectively. For detecting the chromosomes with QTL effects, one can consider for every marker m the statistic χ^2_m taken over all populations (or RIL families) polymorphic for the marker m ,

$$\chi^2_m = \Sigma_f D_{f,m}^2 / \text{Var} D_{f,m}^2, \quad (6.9)$$

where $\text{Var} D_{f,m}$ is the sampling variance of $D_{f,m}$ for family f at marker m . When the selected trait is not affected by the tested chromosome (H_0 hypothesis), χ^2_m is presumed to follow a χ^2 -distribution with d.f. = 1, enabling a χ^2 -test for the presence of a QTL linked to the marker (Weller et al. 1990). The main shortcoming of this approach is that test statistics (9) are considered separately per each marker and information about neighborhoods of significant or non significant markers is not taken into account. This makes difficult to detect whether significance of a given group of markers is caused by its linkage with QTL(s) or only by random sample effect. This problem becomes especially important in situations where individual significance of markers linked to QTL is not extremely high and the total number of markers is high.

By joint analysis of such marker-allele frequency differences, D_m , across marker loci, one can estimate the chromosomal position of the detected QTL. For one or several populations segregating for the same QTL, fitting a function of chromosomal positions for observed D_m values at the polymorphic marker loci can be used for estimation of the QTL position (similar to the procedures described by Kearsey 1998; Ronin et al. 1999; Wang et al. 2007; Korol et al. 2007b). Such approaches are based on idea that the expectation of statistic D_m is proportional to $(1-2r_m)$, where r_m is the effective recombination rate between the marker m and the QTL. Value r_m depends on the location of marker m , unknown location ($x_{(q)}$) of the putative QTL on the chromosome, reproduction type, and population structure. Nevertheless, it is difficult to obtain statistical properties (p -value, confidence intervals for parameter estimates, and QTL detection power). Partially this problem can be solved by applying maximum-likelihood principle (Wang et al. 2007), but it is difficult to find appropriate asymptotic for constructing likelihood function.

Fractionating of pools allows constructing empirical distribution of estimates by conducting the analysis on subsets of data (jackknifing). It also allows constructing empirical distribution of estimates under hypothesis H_0 (no QTLs) by conducting analyses of sets of data obtained by random reshuffling of the subpools between “Low” and “High” tails of trait distribution (FPD permutation test). Comparison of statistic value in initial data configuration with distribution under reshuffling allows estimating directly the significance of QTL effect. Empirical distribution of the estimated parameters on jackknife resampling runs allows constructing confidence interval for these statistics. Comparison of distributions in jackknife runs and permutation runs allow estimating statistical power of QTL detection. Therefore, the proposed FPD

actually extends the methodology developed for individual-based analysis in selective genotyping (e.g., Ronin et al. 1998) to subpool-based analysis in selective DNA pooling.

The number of subpools limits the number of possible different resamplings and permutations and hence limits the level of detectable p-values and resolution of confidence interval. High number of subpools increases the cost of genotyping and hence reduces the advantages of pooling over individual analysis. Mapping resolution can be significantly increased by joint (parallel) analyzing of a set of populations (families) presumably polymorphic for common QTLs. In such cases the number of different possible jackknives and permutations increases polynomially with the number of populations. Fractionating of pools also allows estimating the test statistic more accurately than in the case of a single pool per tail. The procedures considered in Korol et al. 2007b were suitable for mapping populations composed of full- or half-sib families or multiple F_2 or BC families. Here, using simulated data, we illustrate how FPD can be used in linkage analysis (e.g., with RIL) or association (LD) analysis.

6.3.2 Linkage Analysis (RIL)

In the case of RIL data obtained after several generations of selfing, value r_m can be estimated using Haldane-Waddington model with corresponding “map expansion”. Hence the analysis can be conducted by the way analogous to the procedures proposed in Korol et al. 2007b. In the case of one or several QTLs per chromosome, expected D at a marker locus is defined by the frequencies of alleles in the high and low pools at the closest situated QTLs and by recombination rates between marker and QTL. Let K be the number of QTLs in the chromosome and enumeration of the QTLs is according to their locations, i.e., $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(K)}$. Fitting the model for values $D_{f,m}$ observed in population f at markers $1, \dots, M$ can then be written in the matrix form (see also Wang et al. 2007):

$$D_f = X_f \lambda_f + e_f, \quad (6.10)$$

where D_f is a vector of $D_{f,m}$; λ_f is a vector of $\lambda_{f,q}$ characterizing the D -value at putative q th QTL in population f ; X_f is matrix of coefficients defined by mutual arrangement of K QTLs and M markers in the chromosome; e_f is a vector of residuals at marker loci. Criterion for parameter estimation can be formulated in terms of maximal likelihood or least square approach. Variation of variances and correlation between D -values for linked markers also can be taken into account (Korol et al. 2007b).

The confidence intervals for estimated parameters can be obtained using likelihood criteria (Wang et al. 2007) or sampling distribution of estimates of parameters in re-sampling iterations. Testing of hypotheses on QTL-existence and calculation of power of tests can be conducted using resampling analysis

combined with permutation tests (Korol et al. 2007b). The characteristic feature of our FPD is that we do not use likelihood function and, hence, we are free from assumptions about asymptotic normality of the estimates and problematic models for estimation of variances.

Even in the case of only two QTLs on the chromosome, various situations can exist. These include polymorphism of different populations included to the analysis for one, two, or none of the QTLs, and the linkage phase between the QTLs (coupling vs. repulsion) in the populations segregating for both QTLs. Thus, in addition to the tests of significance proposed in Korol et al. 2007b, the situation with linked QTLs calls for comparisons of H_2 vs. H_1 (two-QTL vs. single-QTL hypotheses) for the entire data set as well as for each population.

6.3.3 Association Analysis

The ever growing interest of the cereal mapping community to association analysis for fine genetic mapping justifies extension of the foregoing FPD procedures from linkage analysis to association (LD-based) QTL mapping. In the analysis of real populations detectable LD are usually observed only on very short distances even for such highly selfing plant as wild barley (Flint-Garcia et al. 2003). Hence high density of markers is a critical point here. Still, in other cases selfers may display rather high distances of association (Morrell et al. 2005). Absence of information on characteristics of linkage of QTLs with neighbor markers makes the fitting of corresponding model (curve) to observed data a problem. Individual significance of markers corrected by FDR is also not very informative here, because the number of markers is very high (association preserved in 20 kb, i.e. about 0.05 cM, implies that the number of needed markers is at least 2000). Such situation may be very disappointing when the individual significance of each marker is not extremely high. These problems can be solved by dividing the explored chromosome on overlapping parts. In addition to filtering properties of this approach, the significance of a part (window) containing dozens of markers can be much higher than that of every individual marker.

Let D_I be the mean value of observed D for markers situated on interval I . Within the framework of FPD, chromosome-wise p -value can be obtained for the i th maximum of D_{Ij} by permutation test. Let $D^{(p,i)}$ be the i th maximum of value D_{Ij} in the pool permutated configuration #ps (case $p=0$ corresponds to not-permutated pool configuration). The interval where this maximum is achieved we denote by $I_{(p,i)}$. Then, p -value of the first maximum can be estimated by the proportion of permutations p , where $D^{(p,1)} \geq D^{(0,1)}$; p -value of each next i -th maximum can be estimated by the proportion of permutations p where $D^{(p,i)} \geq D^{(0,i)}$. Here $D^{(p,i)}$ is the i th maximum D_{Ij} on the chromosome with excluded intervals $I_{(0,j)}, j=1, \dots, i-1$. Intervals $I_{(0,j)}$ with significant $D^{(0,j)}$ can be considered as associated with QTL $_j$. More accurate estimation of QTL

position can be obtained by fitting linear regression model within interval $I_{(0,j)}$. Corresponding confidence intervals can be obtained using distribution of estimates in jackknife resampling iterations. The power of detection also can be obtained using jackknife runs and permutations (analogously to Korol et al. 2007b). If distribution of D_I under hypothesis H_0 varies along the chromosome, then some normalization of D_I values is needed.

6.3.4 Simulations

The simulated examples were employed to illustrate the proposed methodology for linkage analysis on RIL and association analysis. Each simulated population (family) consisted of diploid individuals with non-overlapping generations. Individuals were presented by genotypes on $M = 2000$ equidistant di-allelic loci (with alleles “0” and “1”) situated in one chromosome of $L = 100$ cM. Some of these loci were considered as affecting trait value (QTLs), the reminders were considered as SNP markers. Trait values were modeled as normally distributed with variance 1 and mean value equal to sum of additive effects d_{fq} of alleles “1” in QTLs. In the simulations of RIL data all individuals were obtained by selfing, where for association analysis we simulated predominately selfing population with proportion of selfing 0.99. Individuals with extreme trait values (from $\alpha^{(L)}$ and $\alpha^{(H)}$ tails of trait distributions) were randomly united to $n^{(L)}$ and $n^{(H)}$ subpools (bulks). For simplicity, we used $\alpha^{(L)} = \alpha^{(H)} = \alpha$ and $n^{(L)} = n^{(H)} = n$, hence for population of N individuals we had $2n$ subpools of α N/n individuals. Allele frequency in each pool in each marker was noised by random value evenly distributed in interval $[-0.05, 0.05]$ (“error of pooling machine”). In the simulations of RIL data ancestor was with a genotype consisted of “0” and “1” haplotypes. Six generations of selfing were produced for each RIL family. In the simulation of data for association analysis initial linkage disequilibrium (LD) was achieved by simulating initial population of small size with random genotypes. Initial frequencies of QTL-alleles “1” were 0.2. Initial linkages were mostly destroyed during G generations of partial selfing. During this time allele frequencies changed by random drift caused by finite population size; no mutations were assumed.

6.3.5 Example of RIL Data Analysis by FPD

The example presented in Fig. 6.6 illustrates joint fitting of 2-QTL model to RIL data obtained from three simulated families. The first family, **Fam1**, was simulated polymorphic for two QTLs: QTL1 situated in $x_{(1)} = 30$ cM with the allele substitution effect $d_{11} = 0.25$, and QTL2 situated in $x_{(2)} = 70$ cM with the allele substitution effect $d_{12} = 0.2$. The second family, **Fam2**, was polymorphic only for QTL2 with $d_{22} = 0.2$. The third family, **Fam3**, was monomorphic for

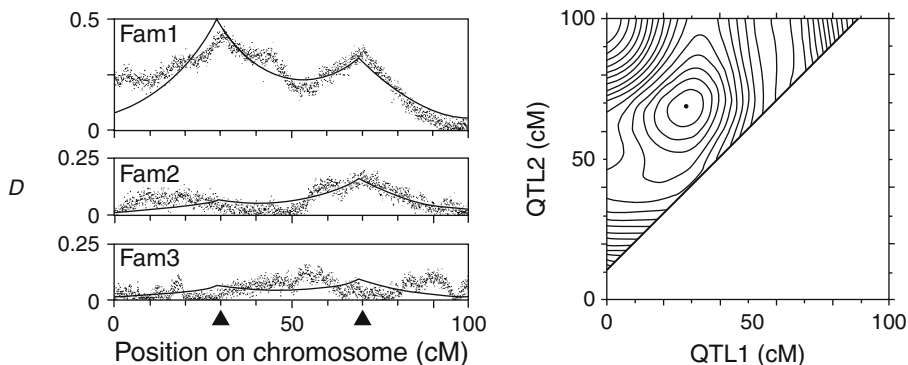


Fig. 6.6 Parallel fitting of 2-QTL model to simulated RIL data. The left graphs demonstrate the model fitting for data on three simulated RIL families. Solid curves correspond to the fitted model. Black triangles mark positions of simulated QTLs. Family **Fam1** was polymorphic for QTL1 and QTL2; **Fam2** was polymorphic only for QTL2; and **Fam3** was monomorphic for both QTLs. The right graph demonstrates the landscape of criterion (10) value for different positions of two QTLs (with restriction on distance between QTLs $x_{(2)} - x_{(1)} \geq 10$ cM). Bold point marks the coordinate of the minimum (i.e., the solution). Contours of levels can be used for constructing of the joint confidence intervals for positions of the QTLs

both QTLs. Every family was of $N=1000$ individuals, tail percentage was $\alpha=0.2$, number of pools per tail was $n=5$ (40 individuals per pool). Positions of two QTLs were estimated as 29 cM and 69 cM. 99% confidence interval for QTL positions obtained by 80% jackknifes were [28, 31] cM for QTL1 and [66, 73] cM for QTL2. There exist 512 different pool permutations for each of the three families. Hence, the total number of different permutations is 512^3 . P -values in intra-family permutation tests were estimated by 0.002 for λ_{11} , 0.008 for λ_{12} , 0.004 for λ_{22} , and more than 0.1 (non-significant) for each of λ_{21} , λ_{31} , and λ_{32} . Combining families with significant effects (**Fam1** and **Fam2**) provide further improvement of significance and mapping accuracy.

6.3.6 Example of Association Analysis by FPD

Example presented in Fig. 6.7 demonstrates simulated data for association analysis. Population of $N=10000$ diploids with $p_{\text{selfing}}=0.99$ was iterated $G=5000$ generations. Two QTLs with equal allele substitution effects $d_1=d_2=0.25$ were simulated in $x_{(1)}=30$ cM and $x_{(2)}=60$ cM. Individuals of the last generation were united in pools and marker allele “1” frequencies were calculated. Tail proportion was $\alpha=0.1$, number of pools per tail was $n=10$ (100 individuals per pool). About 10^5 different pool permutations are possible in this setup. The method described above allows estimating QTL positions and

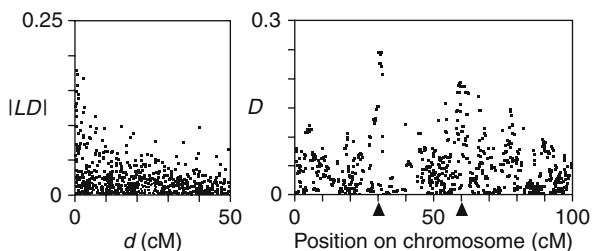


Fig. 6.7 Simulated data for association analysis. The left graph demonstrates decreasing of linkage disequilibria with increasing distance d between loci. The right graph shows D -values for markers. Positions of simulated QTLs are marked by black triangles

p -values: estimated positions for QTL1 and QTL2 were 30.0 and 58.8 cM, respectively. Both of them were with experiment-wise p -value = 0.000011. Additionally to these “true” highly significant QTLs, it was also found three significant ($p < 0.05$) “false” QTLs: at 5.05 cM with $p = 0.003$, at 77.8 cM with $p = 0.003$, and at 66.8 cM with $p = 0.012$. In our analysis we allowed distance between QTLs to be at least 5 cM. The situation can be improved by joint analysis of several populations, because “false” QTLs are expected to arise in different parts of chromosome in different populations (not shown).

6.4 Final Comments

A new approach of multilocus map ordering was presented here that we develop with an objective to increase the reliability of genetic maps. Its characteristic features of is employment of powerful optimization heuristics for reaching the best order and re-sampling based testing of the map quality. Detection and removing markers responsible for local map instabilities and non-monotonic change in recombination rates allows building stable skeleton maps with minimal total length. Further improvement of mapping quality is achievable by joint analysis of mapping data from different mapping populations. Separate ordering of different data sets does not guarantee obtaining identical orders for shared markers in resulting maps calling for detection and removing conflicting markers. An alternative (presented above) is building de novo a consensus multilocus map based on “synchronized ordering” rather than merging the previously derived maps. This approach is also applicable in situations of gender dependent recombination and combined analysis of genetic and physical mapping data, possibly in sequential experimentation manner.

We presented here our “multiple” approach for QTL mapping. It was demonstrated that combined MultipleTrait-Multilocus (MT-MIM) and MultipleEnvironment-Multilocus (ME-MIM) analyzes may serve as useful schemes for improving QTL mapping efficiency, including higher QTL detection power and narrower confidence interval of the estimated QTL position; efficient

analysis of epistasis; detecting QTL-environment interaction; etc. But efficient tools should be used with all precautions of biological justification. It should be noted that multivariate methods do not automatically guarantee an improvement of QTL mapping results (Korol et al. 2001). Indeed, combining in one multivariate analytical framework multiple traits or a trait scored across multiple environments may be associated with both technical obstacles and principal complications. In some situations many QTLs spread along a chromosome may affect the scored traits, resulting in a decrease rather than an increase of mapping precision in joint mapping. Nevertheless, despite these considerations, analysis of numerous data sets points to a tendency of increasing QTL detection power and mapping accuracy (sometimes dramatic) by moving from simplistic (univariate) to more sophisticated (multivariate) methods. Seemingly, there is some dissonance between the ever increasing complexity and cost of generating experimental data and an inclination of a considerable part of experimentalists to analyze the data in an easy, simple and fast way. To our opinion, any complication of mapping models is justified if it increases the quality and amount of information extracted from the generated data.

Acknowledgments This research is partially supported by the Israeli Ministry of Absorption and the United States-Israel Binational Agricultural Research and Development Foundation (grant # 9615).

References

- Ben-Dor, A., Chor, B., and Pelleg, D. (2000). RHO-Radiation hybrid ordering. *Genome Res.* 10: 365–378.
- Boyko, E., Kalendar, R., Korzun, V., Fellers, J., Korol, A., Schulman, A.H., and Gill, B.S. (2002). A high-density cytogenetic map of the *Aegilops tauschii* genome incorporating retrotransposons and defense-related genes: insights into cereal chromosome structure and function. *Plant Mol. Biol.* 48: 767–790.
- Brohede, J., Dunne, R., McKay, J.D., and Hannan, G.N. (2005). PPC: an algorithm for accurate estimation of SNP allele frequencies in small equimolar pools of DNA using data from high density microarrays. *Nucl. Acids Res.* 33: e142.
- Carleos, C., Baro, J.A., Canon, J., and Corral, N. (2003). Asymptotic variances of QTL estimators with selective DNA pooling. *J. Hered* 94: 175–179.
- Churchill, G.A. and Doegre, R.W. (1994). Empirical threshold values for quantitative trait mapping. *Genetics* 138: 963–971.
- Darvasi, A. and Soller, M. (1994). Selective DNA pooling for determination of linkage between a molecular marker and a quantitative trait locus. *Genetics* 138: 1365–1373.
- Dekkers, J.C.M. (2000). Quantitative trait locus mapping based on selective DNA pooling. *Anim. Breed. Genet.* 117: 1–16.
- Denell, R.E. and Keppy, D.O. (1979). The nature of genetic recombination near the third chromosome centromere of *Drosophila melanogaster*. *Genetics* 93: 117–130.
- Dobzhansky, T.H., Spassky, B., and Anderson, W. (1965). Bichromosomal synthetic semi-lethals in *Drosophila pseudoobscura*. *Proc. Nat. Acad. Sci. USA* 53: 345–348.
- Eberhard, S.A. and Russel, W.A. (1966). Stability parameters for comparing varieties. *Crop Sci.* 6: 36–40.
- Efron, B. (1979). Bootstrap method: another look at the jackknife. *Ann. Stat.* 7: 1–26.

- Emrich, S.J., Aluru, S., Fu, Y., Wen, T.J., Narayanan, M., Guo, L., Ashlock, D.A., and Schnable, P.S. (2004). A strategy for assembling the masize (*Zea mays* L.) genome. *Bioinformatics* 20: 140–147.
- Esch, E. and Weber, W.E. (2002). Investigation of crossover interference in barley (*Hordeum vulgare* L.) using the coefficient of coincidence. *Theor. Appl. Genet.* 104: 786–796.
- Falk, C.T. (1992). Preliminary ordering of multiple linked loci using pairwise linkage data. *Genet. Epidemiol.* 9: 367–375.
- Finlay, K.W. and Wilkinson, G.N. (1963). The analysis of adaptation in a plant-breeding programme. *Aust. J. Agric. Res.* 14: 742–754.
- Flint-Garcia, S.A., Thornsberry, J.M., and Buckler, E.S. (2003). Structure of linkage disequilibrium in plants. *Annu. Rev. Plant Biol.* 54: 357–374.
- Givry, S., Bouchez, M., Chabrier, P., Milan, D., and Schiex, T. (2005). CarthaGene: multi-population integrated genetic and radiation hybrid mapping. *Bioinformatics* 8: 1703–1704.
- Hayes, P.M., Liu, B.H., Knapp, S.J., Chen, F., Jones, B., Blake, T., Franckowiak, J., Rasmuson, D., Sorrels, M., Ullrich, S.E., Wesenberg, D., and Kleinhofs, A. (1993). Quantitative trait locus effects and environmental interaction in a sample of North American barley germplasm. *Theor. Appl. Genet.* 87: 392–401.
- Hillel, J., Avner, R., Baxter-Jones, C., Dunnington, E.A., Cahaner, A. et al. (1990). DNA fingerprints from blood mixes in chickens and turkeys. *Anim. Biotechnol.* 2: 201–204.
- Jansen, R.C. and Stam, P. (1994). High resolution of quantitative t1r4a5it5s. into multiple loci via interval mapping. *Genetics* 136: 1447–1455.
- Jansen, J., de Jong, A.G., and Ooijen, J.W. (2001). Constructing dense genetic linkage maps. *Theor. Appl. Genet.* 102: 1113–1122.
- Jansen, R.C., Van Ooijen, J.M., Stam, P., Lister, C., and Dean, C. (1995). Genotype-by-environment interaction in genetic mapping of multiple quantitative trait loci. *Theor. Appl. Genet.* 91: 33–37.
- Jiang, C. and Zeng, Z.-B. (1995). Multiple trait analysis and genetic mapping for quantitative trait loci. *Genetics* 140: 1111–1127.
- Johnson, T. (2005). Multipoint linkage disequilibrium mapping using multilocus allele frequency data. *Ann. Hum. Genet.* 69: 474–497.
- Joppa, L.R., Nevo, E., and Beiles, A. (1995). Chromosome translocations in wild populations of tetraploid emmer wheat in Israel and Turkey. *Theor. Appl. Genet.* 91: 713–719.
- Kao, C.-H., Zeng, Z.-B., and Teasdale, R.D. (1999). Multiple interval mapping for quantitative trait loci. *Genetics* 152: 1203–1216.
- Kearsey, M.J. (1998). The principles of QTL analysis (a minimal mathematics approach). *J. Exp. Bot.* 49: 1619–1623.
- Klein, P.E., Klein, R.R., Cartinhour, S.W., Ulanich, P.E., Dong, J., et al. (2000). A High-throughput AFLP-based method for constructing integrated genetic and physical maps: progress toward a sorghum genome map. *Genome Res.* 10: 789–807.
- Korol, A.B., Preygel, I.A., and Bocharnikova, N.I. (1987). Linkage between loci of quantitative characters and marker loci. 5. Combined analysis of several markers and quantitative characters. *Genetika (USSR)* 23: 1421–1431 (in Russian). English translation in *Soviet Genetics* 1988, 23: 996–1004 (Plenum Publ. Co., N.Y.).
- Korol, A.B., Preygel, I.A., and Preygel, S.I. (1994). *Recombination Variability and Evolution*. Chapman & Hall, London.
- Korol, A.B., Ronin, Y.I., and Kirzhner, V.M. (1995). Multitrait analysis in interval mapping of QTL. *Genetics* 140: 1137–1147.
- Korol, A., Ronin, Y., Tadmor, Y., Bar-Zur, A., Kirzhner, V.M., and Nevo, E. (1996). Estimating variance effect of QTL: An important prospect to increase the resolution power of interval mapping. *Genet. Res.* 67: 187–194.
- Korol, A.B., Ronin, Y.I., and Nevo, E. (1998). Approximated analysis of QTL-environmental interaction with no limits on the number of environments. *Genetics* 148: 2015–2028.

- Korol, A., Ronin, Y., Itzcovich, A., and Nevo, E. (2001). Enhanced efficiency of QTL mapping analysis based on multivariate complexes of quantitative traits. *Genetics* 157: 1789–1803.
- Korol, A., Shirak, A., Cnaani, A., and Hallerman, E.M. (2007a). Detection and analysis of QTLs for economic traits in aquatic species. In: Liu, Z.J. (ed.), *Aquaculture Genome Technologies*. Blackwell, pp. 169–197.
- Korol, A., Frenkel, Z., Cohen, L., Lipkin, E., and Soller, M. (2007b). Fractioned DNA Pooling: A New Cost-Effective Strategy for Fine Mapping of Quantitative Trait Loci. *Genetics* 176: 2611–2623.
- Korzun, V., Malyshev, S., Voylovkov, A.V., and Börner, A. (2001). A genetic map of rye (*Secale cereale* L.) combining RFLP, isozyme, protein, microsatellite and gene loci. *Theor. Appl. Genet.* 102:709–717.
- Lacaze, X., Tanny, S., and Korol, A. (2009a). Transcriptional plasticity differing across genetic backgrounds: An epistatic mechanism in *Arabidopsis thaliana* (in revision).
- Lacaze, X., Hayes, P.M., and Korol, A. (2009b). Genetics of phenotypic plasticity: QTL analysis in barley, *Hordeum vulgare*. *Heredity* 102:163–173.
- Lander, E.S. and Botstein, D. (1989) Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* 121, 185–199.
- Liu, B.H. (1998). *Statistical Genomics: Linkage, Mapping, and QTL Analysis*. CRC Press, New York.
- Liu, C.J., Devos, K.M., Chinoy, C.N., Atkinson, M.D., and Gale, M.D. (1992). Non-homoeologous translocations between group-4, 5 and 7 chromosomes in wheat and rye. *Theor. Appl. Genet.* 83: 305–312.
- Mangin, B., Thoquet, P., and Grimsley, N. (1998). Pleiotropic QTL analysis. *Biometrics* 54: 88–99.
- Menotti-Raymond, M., David, V.A., Chen, Z.Q., Menotti, K.A., Sun, S., Schaffer, A.A., Agarwala, R., Tomlin, J.F., O'Brien, S.J., and Murphy, W.J. (2003). Second-generation integrated genetic linkage/ radiation hybrid maps of the domestic cat (*Felis catus*). *J. Hered.* 94: 95–106.
- Mester, D., Ronin, Y., Minkov, D., Nevo, E., and Korol, A. (2003a). Constructing large scale genetic maps using Evolutionary Strategy Algorithm. *Genetics* 165: 2269–2282.
- Mester, D., Ronin, Y., Hu, Y., Nevo, E. and Korol, A. (2003b). Efficient multipoint mapping: Making use of dominant repulsion-phase markers. *Theor. Appl. Genet.* 107, 1102–1112.
- Mester, D.I., Ronin, Y.I., Nevo, E., and Korol, A.B. (2004). Fast and high precision algorithms for optimization in large scale genomic problems. *Comp Biol & Chemistry* 28: 281–290.
- Mester, D., Ronin, Y., Korostishevsky, M., Pikus, V., Glazman, A., and Korol, A.B. (2005). Multilocus consensus genetic maps (MCGM): Formulation, algorithms and results. *Computat. Biol. Chem.* 30: 12–20.
- Michie, D. (1953). Affinity: a new genetic phenomenon in the house mouse. *Nature* 171: 26–27.
- Morrell, P.L., Toleno, D.M., Lundy, K.E., and Clegg, M.T. (2005). Low levels of linkage disequilibrium in wild barley (*Hordeum vulgare* ssp *spontaneum*) despite high rates of self-fertilization. *Proc. Natl. Acad. Sci. USA* 102: 2442–2447.
- Morrell, P.L., Toleno, D.M., Lundy, K.E., and Clegg, M.T. (2006). Estimating the contribution of mutation, recombination and gene conversion in the generation of haplotypic diversity. *Genetics* 173: 1705–1723.
- Peng, J., Korol, A.B., Fahima, T., Roder, M.S., Ronin, Y.I., Li, Y.C., and Nevo, E. (2000). Molecular genetic maps in wild emmer wheat, *Triticum dicoccoides*: genome-wide coverage, massive negative interference, and putative quasi-linkage. *Genome Res.* 10: 1509–1531.
- Peng, J., Ronin, Y., Fahima, T., Röder, M.S., Li, Y., Nevo, E., and Korol, A.B. (2003). Domestication quantitative trait loci in *Triticum dicoccoides*, the progenitor of wheat. *Proc. Natl. Acad. Sci. USA* 100: 2489–2494.

- Plagnol, V., Padhukasahasram, B., Wall, J.D., Marjoram, P., and Nordborg, M. (2006). Relative influences of crossing over and gene conversion on the pattern of linkage disequilibrium in *Arabidopsis thaliana*. *Genetics* 172: 2441–2448.
- Romagosa, I., Ullrich, S.E., Han, F., and Hayes, P.M. (1996). Use of additive main effects and multiplicative interaction model in QTL mapping for adaptation in barley. *Theor. Appl. Genet.* 93: 30–37.
- Ronin, Y.I., Kirzhner, V.M., and Korol, A.B. (1995). Linkage between loci of quantitative traits and marker loci. Multitrait analysis with a single marker. *Theor. Appl. Genet.* 90: 776–786.
- Ronin, Y.I., Korol, A.B., and Weller, J.I. (1998). Selective genotyping to detect quantitative trait loci affecting multiple traits: interval mapping analysis. *Theor. Appl. Genet.* 97: 1169–1178.
- Ronin, Y., Korol, A., and Nevo, E. (1999). Single- and multiple-trait analysis of linked QTLs: some asymptotic analytical approximation. *Genetics* 151: 387–396.
- Sari-Gorla, M., Calinski, T., Kaczmarek, Z., and Krajewski, P. (1997). Detection of QTL-environment interaction in maize by a least squares interval mapping method. *Heredity* 78: 146–157.
- Schiex, T. and Gaspin, C. (1997). Carthagine: constructing and joining maximum likelihood genetic maps. *ISMB* 5: 258–267.
- Sinclair, D.A. (1975). Crossing over between closely linked markers spanning the centromere of chromosome 3 in *Drosophila melanogaster*. *Genet. Res.* 11: 173–185.
- SØgaard, B. (1977). The localization of *eceriferum* loci in barley. V. Three point tests of genes on chromosome 1 and 3 in barley. *Carlsberg Res. Commun.* 42: 67–75.
- Voudouris, C. (1997). *Guided local search for combinatorial problems*, Ph.D. thesis, Department of Computer Science, University of Essex, Colchester.
- Wang, J., Koehler, K.J., and Dekkers, J.C.M. (2007). Interval mapping of quantitative trait loci with selective DNA pooling data. *Genet. Select. Evol.* 39: 685–710.
- Weeks, D. and Lange, K. (1987). Preliminary ranking procedures for multilocus ordering. *Genomics* 1: 236–242.
- Weller, J.I., Kashi, Y., and Soller, M. (1990). Power of daughter and granddaughter designs for determining linkage between marker loci and quantitative trait loci in dairy-cattle. *J. Dairy Sci.* 73: 2525–2537.
- Weller, J.I., Wiggans, G.R., Van Raden, P.M., and Ron, M. (1996). Application of a canonical transformation to detection of quantitative trait loci with the aid of genetic markers in a multi-trait experiment. *Theor. Appl. Genet.* 92: 998–1002.
- West, M.A.L., Kim, K., Kliebenstein, D.J., van Leeuwen, H., Michelmore, R.W., Doerge, R.W., and St. Clair, D.A. (2007). Global eQTL mapping reveals the complex genetic architecture of transcript-level variation in *Arabidopsis*. *Genetics* 175: 1441–1450.
- Yap, I., Schneider, D., Kleinberg, J., Matthews, D., Cartinhour, S., and McCouch, S. (2003). A graph-theoretic approach to comparing and integrating genetic, physical and sequence-based maps. *Genetics* 165: 2235–2247.
- Zeng, Z.-B. (1994). Precision mapping of quantitative trait loci. *Genetics* 136: 1457–1468.

Chapter 7

Genetic Mapping in the Triticeae

Anke Lehmensiek, William Bovill, Peter Wenzl, Peter Langridge,
and Rudi Appels

Abstract Genetic maps are the fundamental tools to identify features of phenotypes that are linked to specific genetic loci and eventually DNA sequences or genes. The major use of genetic linkage maps has, therefore, been to identify quantitative trait loci (QTL). Genetic maps are also essential for marker-assisted selection, comparative mapping, high-resolution mapping and map-based cloning. To date, over 40 maps with at least 300 markers have been published for different Triticeae populations. The quality of genetic maps can be affected by a number of factors and map curation ensures that map quality issues are identified and, where possible, resolved. We report on the issues involved in the production of quality genetic linkage maps by inspection of marker genotype data after map construction.

A number of technologies, which have been developed to complement the genetic linkage maps of the Triticeae such as radiation induced deletion mutations, optical mapping and HAPPY mapping, are discussed. We also report on the construction of consensus maps and the issues involved in the building of these.

Genetic map quality, including the ordering of loci within linkage groups, is of great importance for robust QTL detection and a number of QTL mapping methods are discussed in relation to features of the genetic map used in the analysis. Finally, we describe high-resolution mapping, which is used to improve the confidence interval for a QTL and to obtain markers closer to the trait of interest.

A. Lehmensiek (✉)

Centre for Systems Biology, University of Southern Queensland, Toowoomba,
Australia

e-mail: lehmentsie@usq.edu.au

7.1 Introduction

The Triticeae tribe includes a broad range of species, some of which (*Triticum aestivum*, *Hordeum vulgare*, and *Secale cereale* in particular) dominate agricultural landscapes in temperate areas. Apart from the value of their grain for human consumption and animal feed, some species (e.g. *Elymus* and *Agropyron*) are also important forage grasses. Research into the Triticeae is crucial in a world where human populations are expanding and where undeveloped arable land is becoming increasingly scarce. The production of new high-yielding varieties that are adapted to abiotic and biotic stresses is laborious, time-consuming, and reliant on careful phenotypic selection (Kumar 1999). Through conventional breeding strategies, the production of a new wheat variety for example, can take up to 14 years. The pressure to produce new varieties more rapidly is increasing. Indeed, in 2006, global consumption of wheat was forecast to exceed supply for the second year running, with ending stocks forecast to reach their lowest level in 25 years (United States Department of Agriculture 2006).

Biotechnology offers plant breeders new methods that have the potential to increase the rate at which new varieties are bred and to improve the characteristics that are desired for their end-use. Molecular marker technology is a powerful tool that can be used to increase the understanding and subsequent manipulation of the genetics of both simple and complex traits (Dubcovsky 2004). With simple traits, markers tightly linked to the gene of interest can be used to indirectly select for the desirable allele (Anderson et al. 1989; Peng et al. 1999; Beecher et al. 2002). For more complex traits, genetic linkage maps composed of molecular markers can be used to identify regions of the genome that contribute to phenotypic variation (Parker et al. 1999; Chartrain et al. 2004; Chen et al. 2006). Markers linked to quantitative trait loci (QTL) can be used to more rapidly incorporate desirable regions into agronomically superior genotypes. The purpose of this review is to focus on genetic mapping in the Triticeae, with an emphasis on the production of high quality genetic linkage maps for QTL mapping and map-based cloning of genes. Only the economically important Triticeae species (*Triticum aestivum*, *Hordeum vulgare*, *Secale cereale*, and \times *Triticosecale*), for which a large number of genetic mapping studies have been reported, will be discussed.

Wheat (*Triticum* spp.) is one of the three most important cereal crops for humankind – in the form of bread, wheat provides more nutrients to the world population than any other single food source (Pena 2002). Wheat has three levels of ploidy (number of copies of the basic number of chromosomes): diploid ($2n = 2 \times = 14$), tetraploid ($2n = 4 \times = 28$) and hexaploid ($2n = 6 \times = 42$). Of the three ploidy levels, tetraploid (durum) and hexaploid (bread) wheats are the most commonly grown wheats in present-day agriculture. Cultivated barley (*Hordeum vulgare*) is diploid ($2n = 2 \times = 14$), although tetraploid and hexaploid wild relatives exist (Martin et al. 1999). Barley is grown in a range of extreme environments, generally for malting purposes and animal feed. The genus *Hordeum* comprises 32

species, of which *Hordeum vulgare* is the major domesticated species. It was derived from the wild progenitor *Hordeum spontaneum*, which continues to grow as a wild relative in the Middle East. Barley can be divided by the number of kernel rows in the head. Three forms have been cultivated; two-row barley (*Hordeum distichum* and *Hordeum vulgare*), four-row barley (*Hordeum tetrastichum*) and six-row barley (*Hordeum vulgare*; Badr et al. 2000). Rye (*Secale* spp.) is a diploid ($2n = 2 \times = 14$), grown as a grain and forage crop. The genus *Secale* is composed of four species – *S. cereale* L. (outbreeding annual), *S. sylvestre* Host (inbreeding annual), *S. vavilovii* Grossh (inbreeding annual), and *S. strictum* (Presl.) Presl. (syn. *S. montanum*; outbreeding perennial) (De Bustos and Jouve 2002). Subspecies of *S. cereale* and *S. strictum* exist, although the only species presently grown in cultivation is *S. cereale* ssp. *cereale* (Shang et al. 2006). Triticale (\times *Triticosecale* Wittmack) is a polyploid synthetic cereal species, obtained from crosses between tetraploid or hexaploid wheats and rye ($2n = 6 \times = 42$ or $2n = 8 \times = 56$). Interest in cultivation of triticale continues to be maintained as it provides good yields, performs better than wheat in poor, acidic soils, and is only mildly susceptible to diseases that affect other cereals (González et al. 2005).

Due to their allopolyploid nature, wheat and triticale are the most complex species. The inter-relatedness of the genomes is emphasized by both the man-made formation of triticale where in the case of the hexaploid \times *Triticosecale*, the R genome substitutes for the D genome as well as the range of chromosome addition and substitution lines produced using hexaploid wheat (Sears 1954; Islam et al. 1981; Bonafede et al. 2006; Szakacs and Molnar-Lang 2007). As a result of this inter-relatedness, mapping of EST sequences (for example in wheat) has been used to estimate their physical distribution in the other species (Qi et al. 2003; Francki and Appels 2007). The ability to cross-reference the locations of EST sequences has been particularly useful for mapping in the Triticeae because microsatellites/single sequence repeat (SSR) molecular markers from ESTs have been found to be transferable between species (Aishwarya and Sharma 2007). The *Brachypodium* species also has proven to be a useful genome for cross-referencing to regions of interest in wheat such as the Lr34 locus on chromosome 7D (Schnurbusch et al. 2004 and Spielmeier et al. 2008).

7.2 Genetic Linkage Maps

The Triticeae genomes are characterised by regions of so-called gene islands interspersed by large regions of repetitive DNA. At the macro-level in wheat, for example, Erayman et al. (2004) found that, based upon a sample of 3,025 gene loci, 29% of the wheat genome contains 94% of the genes, with 60% of those genes concentrated in only 11% of the genome. At this macro-level La Rota and Sorrells (2004) have demonstrated that an extremely useful synteny exists between the order of blocks of genes in wheat and *Oryza sativa* (rice). Micro-level comparative genomics sequence analyses of wheat, barley and rice

(Appels et al. 2003), indicates that inversions, deletions, duplications and expansion of the genome between orthologous genes is extensive especially in wheat relative to rice. This expansion of the genome resulting from retrotransposon insertion activity at some stage in the evolution of the Triticeae means that the 'gene island' concept may be difficult to relate to any functional attributes of genome structure.

Regardless of size, the molecular unravelling of the wheat genome has been further confounded by its composition. Wheat is an allopolyploid, formed by the hybridization and subsequent chromosome-doubling of two (durum wheat) or three (bread wheat) diploid donors. The genetic similarity between homoeologous chromosomes in wheat provides a buffering capacity that allows chromosome manipulation to be carried out. In his pioneering work Sears (1954) exploited this synteny to develop a set of aneuploid lines in which homoeologous chromosomes compensate for the absence of others. For example, the line nullisomic 5A/tetrasomic 5B has lost both copies of 5A but has four copies of 5B. Sears and Sears (1978) produced ditelosomic lines that lacked entire chromosome arms. Endo and Gill (1996) characterised a set of deletion lines, in which segments of individual chromosome are missing. The aneuploid (nulli-tetrasomic and deletion) lines have been used to identify the chromosomal location of genes and markers (Gill et al. 1993; Huang et al. 2000; Qi et al. 2003), and have become powerful tools for the detailed genetic mapping of genes and molecular markers in the Triticeae.

Molecular markers are used in the construction of genetic linkage maps. Genetic maps indicate the position and distance between markers, as determined by the frequency of recombination events or crossovers between markers, thereby giving a graphical representation of the arrangement of markers along chromosomes (Collard et al. 2005). Recombination is less likely to occur between markers which reside close together as opposed to markers which are further apart. Distances between markers are measured in centimorgans (cM) where 1 cM is approximately equal to a 1 percent chance that two markers will be separated by recombination during meiosis. For plant breeding programs, the most important use of linkage maps is the identification of regions of the genome which contribute to a phenotype. Four major steps are involved in the construction of a genetic linkage map: (1) production of a mapping population; (2) polymorphism assessment; (3) genotyping the population with polymorphic markers and (4) linkage analysis.

When a whole-genome linkage map of a population is to be produced, it is critical that the parents chosen exhibit sufficient genetic polymorphism and lack co-ancestral genome regions (Young 1996). Although it is possible to undertake a "molecular screen" of the parental genotypes to confirm that sufficient polymorphism exists, more commonly, parent lines are selected based upon how much they differ for the phenotype of interest. The parents are then crossed to produce a segregating population such as an F₂ population; backcross population; recombinant inbred (RI) population; or doubled haploid (DH) population (as examples). Each population type has advantages and disadvantages; the most commonly used population type for genetic mapping in the Triticeae are RI and DH

populations. The advantage of these population types is their fixed nature. This allows populations to be phenotyped in diverse environments and for many traits.

Linkage analysis is conducted on the genotypic data that is produced by screening polymorphic markers across the population of segregating progeny. While it is possible to manually determine position and distance between markers for a small number of markers, the large number of markers used to create linkage maps renders the use of computer programs a necessity. A suite of programs can perform linkage analysis, and these include for example: MapMaker/EXP (Lander et al. 1987); JoinMap (Stam 1993); MapManager QTX (Manly et al. 2001); Multipoint (Mester et al. 2003a,b); CarteBlanche (Keygene, Wageningen, The Netherlands) and RECORD (Van Os et al. 2005a) for marker ordering. With the exception of JoinMap, MultiPoint and Carte-Blanche these computer programs are available freely over the internet.

Over 40 maps with at least 300 markers have been published for different Triticeae populations (Table 7.1). Early genetic linkage maps were composed of restriction fragment length polymorphism (RFLP) markers (Chao et al. 1989; Devos et al. 1992; Graner et al. 1991; Nelson et al. 1995); but with the advent of the PCR-based techniques, these were superseded by maps composed of RAPD (Devos and Gale 1992; Williams et al. 1990), amplified fragment length polymorphism (AFLP) (Vos et al. 1995), SSR (Gupta et al. 2002; Röder et al. 1998; Somers et al. 2004; Varshney et al. 2007), single nucleotide polymorphism (SNP) (Gupta et al. 2001; Rostoks et al. 2005b) and Diversity arrays technology (DArT) markers (Jaccoud et al. 2001). For further discussion of marker technologies see Chapter 9 (Paux and Sourdille).

SNPs are single base-pair changes that are more abundant than polymorphisms in other molecular markers and can occur in both coding and non-coding regions of the genome. SNPs can be deployed in routine marker applications (for example see Li et al. 2008), and, due to their abundance, major wheat and barley projects are currently mapping SNPs (<http://rye.pw.usda.gov/snpworld/Search>; <http://harvest.ucr.edu>). It is envisaged that this marker type will dominate wheat and barley genetics studies in the future.

New marker technologies, such as DArT markers, provide the opportunity to quickly develop genetic maps. DArT is a microarray platform that assays a large number of SNP and insertion/deletion (InDel) polymorphisms at restriction enzyme (RE) sites distributed throughout the genome. Polymorphism detection is based on measuring the presence vs. absence of individual fragments in genome representations prepared by RE digestion and adapter ligation (Kilian et al. 2005). While there are many other possible implementations, the most common format of DArT makes use of the methylation-filtration effect of *Pst*I to enrich representations for hypomethylated, low-copy genome regions which tend to contain actively expressed genes (Wenzl et al. 2004 and Kilian et al. 2005). A sequence analysis of approximately 2,500 DArT markers from a *Pst*I representation of barley, for example, suggested that approximately 50% of the markers were derived from actively expressed sequences (Triticarte P/L, Canberra, Australia). DArT markers are usually assayed in a dominant

Table 7.1 Triticeae genetic linkage maps with more than 300 markers

Species	Map Name	Population size	Number of markers	References
Wheat	Arina × Forno	240	380	Paillard et al. 2003
Wheat	CD87 × Katepwa	180	449	Chalmers et al. 2001; Lehmensiek et al. 2005
Wheat	Chinese Spring × SQ1	96	567	Quarrie et al. 2005
Wheat	Courtot × Chinese Spring	187	659	Sourdille et al. 2003
Wheat	Cranbrook × Halberd	161	1325	Chalmers et al. 2001; Lehmensiek et al. 2005; McFadden et al. 2006; Akbari et al. 2006
Wheat	Grandin × BR34	118	743	Liu et al. 2005
Wheat	RL4452 × ‘AC Domain’	182	369	McCartney et al. 2005; Somers et al. 2004
Wheat	W7984 × Opata	115	>2150	Mingeot and Jacquemin 1999; Song et al. 2005; Gupta et al. 2002; Hayden et al. 2006; Somers et al. 2004
Wheat	Sunco × Tasman	180	361	Chalmers et al. 2001; Lehmensiek et al. 2005
Wheat	WL711 × PH132	100	509	Prasad et al. 2003
Wheat	Wuhan-1 × Maringa	101	600	Somers et al. 2003; Somers et al. 2004
Wheat	Kitamoe × Münstertaler	96	464	Torada et al. 2006
Durum	Colloseo × Lloyd	184	633	Wenzl et al. 2007b
Durum	Jennah Khetifa × Cham1	110	470	Nachit et al. 2001
Durum	Messapia × MG4343	65	315	Korzun et al. 1999
Barley	Alexis × Sloop DH	110	358	Willsmore et al. 2006
Barley	Arapiles × Franklin	225	323	Raman et al. 2003
Barley	Azumamugi × Kanto NakateGold	99	307	Mano et al. 2001
Barley	Barque-73 × CP171284-48	90	1000	Wenzl et al. 2006; Hearnden et al. 2007
Barley	Cebada Capa × SusPtrit	113	496	Marcel et al. 2007
Barley	Chebec × Harrington	112	481	Willsmore et al. 2006
Barley	Clipper × Sahara	146	901	Karkousis et al. 2003; Wenzl et al. 2006; Willsmore et al. 2006
Barley	Dayton × Zhepi2	85	576	Wenzl et al. 2006
Barley	Foster × Clho4196	144	556	Horsley et al. 2006; Wenzl et al. 2006
Barley	Frederickson × Stander	130	523	Mesfin et al. 2003; Wenzl et al. 2006
Barley	Galleon × Haruna Nijo	112	481	Willsmore et al. 2006
Barley	Harbin 2-row × Turkey 6	235	383	Hori et al. 2006
Barley	Igri × Atlas68	54	480	Wenzl et al. 2006
Barley	Igri × Franka	71	609	Graner et al. 1991; Varshney et al. 2006; Varshney et al. 2007; Stein et al. 2007

Table 7.1 (continued)

Species	Map Name	Population size	Number of markers	References
Barley	L94 × Vada	103	958	Qi et al. 1998; Marcel et al. 2007; Varshney et al. 2007
Barley	Lina × Canada Park (H. Spontaneum HS92)	86	424	Ramsay et al. 2000; Varshney et al. 2007; Rostoks et al. 2005a
Barley	Oregon Wolfe Dominant × Oregon Wolfe Recessive	94	1237	Costa et al. 2001; Varshney et al. 2006; Varshney et al. 2007; Stein et al. 2007; Rostoks et al. 2005a
Barley	Patty × Tallon	97	423	Wenzl et al. 2006
Barley	Proctor × Nudinka	113	665	Heun et al. 1991; Castiglioni et al. 1998
Barley	Russia 6 × H.E.S.	95	1782	Hori et al. 2003
Barley	SusPtrit × Vada	152	457	Jafary et al. 2006a; Marcel et al. 2007; Varshney et al. 2007
Barley	SusPtrit × Cebada	113	495	Jafary et al. 2006b; Marcel et al. 2007
Barley	Steptoe × Morex	150	1657	Kleinhofs et al. 1993; Varshney et al. 2006; Wenzl et al. 2006; Varshney et al. 2007; Stein et al. 2007; Rostoks et al. 2005a
Barley	TX9425 × Franklin	89	394	Li et al. 2005, 2006; Wenzl et al. 2006
Barley	Yerong × Franklin	180	477	Wenzl et al. 2006
Rye	DS2 × RXL10	101	480	Devos et al. 1993; Masojć et al. 2001; Bednarek et al. 2003; Börner and Korzun 1998
Triticale	Torote × Presto	73	356	González et al. 2005

fashion. However, a genomic representation typically contains a set of markers that may provide sufficient resolution to identify heterozygotes by carefully quantifying hybridization intensities (Diversity Arrays Technology P/L, Canberra, Australia). The quantitative nature of DArT assays was also evident in a Bulk Segregant Analysis (BSA) experiment; the allele-frequency difference between the bulks was tightly correlated with the measured hybridization difference ($r = 0.96$) (Wenzl et al. 2007a). The hybridization intensities, therefore, are proportional to the abundance of alleles in DNA pools. To this date, more than 120 wheat, 60 barley and several rye mapping populations have been genotyped with DArT (Triticarte P/L and Diversity Arrays Technology P/L, Canberra, Australia). Overall, it is evident that DArT markers are well-suited for rapid linkage mapping, and it seems likely that the number of Triticeae maps in the literature is going to increase considerably in the near future.

7.2.1 Wheat Genetic Linkage Maps

Twelve wheat maps with more than 300 markers are listed in Table 7.1. The populations used to construct these maps contained between 96 and 240 individuals. The resulting maps consist of between 361 and more than 2,150 markers and vary considerably in length. Chalmers et al. (2001) reported a Cranbrook \times Halberd DH map spanning 4,110 cM, a CD87 \times Katepwa DH map spanning 3,484 cM, and a Sunco \times Tasman DH map with 3,164 cM. Map curation, however, reduced the lengths of these maps to 3,072, 2,822 and 2,150 cM, respectively (Lehmensiek et al. 2005). Paillard et al. (2003) reported a distance of 3,086 cM for a linkage map based upon a cross between the two Swiss winter wheat varieties Arina and Forno. Liu et al. (2005) reported a similar distance of 3,045 cM for the Grandin \times Br34 map. The map generated from a cross between cv. Courtot and Chinese Spring by Sourdille et al. (2003) spans 3,685 cM. Somers et al. (2004) joined data from four independent genetic maps to produce a 1,235 SSR consensus map spanning 2,569 cM. Appels (2003) produced a wheat consensus map consisting of 4,030 SSR, RFLP and AFLP markers with a total map distance of 3,235 cM. While the genetic distance of the maps varies considerably, the common ground between these maps is that they are all composed of large numbers of markers. This is highly desirable, because the construction of a genetic linkage map is often the first step to identifying regions of the wheat genome that contribute to the expression of quantitative traits.

The deployment of Multi-parent Advanced Generation Intercross (MAGIC) populations for a large-scale gene-trait analysis has been investigated as a means of establishing a resource of interest to a wide-range of research groups (Cavannagh et al. 2008). Incorporating multiple parents provides for the sampling of greater genetic variation in a structured population for use of both linkage and association methodologies. The recovery of 5000–7000 recombinant inbred lines allows for the modelling of multiple QTL and cytoplasmic effects on traits assayed in the population. The theoretical basis for analyzing marker and trait data for a complex cross system has been well characterized in mammalian systems (Teuscher and Broman 2007).

7.2.2 Durum Genetic Linkage Maps

Three durum maps with more than 300 markers have been constructed (Table 7.1). The maps, Colloseo \times Loyd, Jennah Khetifa \times Cham1 and Messapia \times MG4343, consist of 633, 470 and 315 markers, respectively (Wenzl et al. 2007b; Nachit et al. 2001; Korzun et al. 1999).

7.2.3 Barley Genetic Linkage Maps

The 23 barley maps listed in Table 7.1 were produced by using populations of between 54 and 152 individuals. The number of markers mapped on these

populations ranged from 307 to 1,000, and the total map lengths ranged from 926 cM for the Azumamugi \times Kanto Nakate Gold map (Mano et al. 2001) to 1,596 cM for the Russia 6 \times H.E.S. map (Hori et al. 2003). Five consensus maps have been produced in barley. The three DH populations, Steptoe \times Morex, Oregon Wolfe Dominant \times Oregon Wolfe Recessive and Lina \times Canada Park were used to construct an integrated SNP map incorporating 1,237 loci and covering 1,211 cM (Rostoks et al. 2005a). Marcel et al. (2007) produced a consensus map from 3 recombinant inbred line (RIL) populations (L94 \times Vada, SusPtrit \times Vada and SusPtrit \times Cebada Capa) and 2 DH populations (Steptoe \times Morex and Oregon Wolfe Dominant \times Oregon Wolfe Recessive). The final map consisted of 3,258 markers and the total map distance was 1,081 cM. Six independent genetic maps based on the populations Igri \times Franka, Steptoe \times Morex, Oregon Wolfe Dominant \times Oregon Wolfe Recessive, Lina \times Canada Park, L94 \times Vada and SusPtrit \times Vada were used by Varshney et al. (2007) to produce a 775-locus consensus map spanning 1,068 cM. The consensus map produced by Stein et al. (2007) had a map distance of 1,118 cM and contained 1,055 loci, including 607 RFLPs, 190 SSRs and 258 SNPs. This map was composed of the 3 DH populations Igri \times Franka, Steptoe \times Morex and Oregon Wolfe Dominant \times Oregon Wolfe Recessive. Wenzl et al. (2006) constructed a consensus map using the 7 DH populations Barque-73 \times CPI71284-48, Clipper \times Sahara, Dayton \times Zhepi2, Igri \times Atlas68a, Steptoe \times Morex, TX9425 \times Franklin and Yerong \times Franklin, and the 3 RI populations Foster \times CI4196 F8-9, Frederickson \times Stander and Patty \times Tallon. The map contained 2,935 markers and spanned 1,161 cM.

7.2.4 Rye Genetic Linkage Maps

To date only one rye map with more than 300 markers has been produced (Table 7.1). This map, constructed for the DS2 \times RXL10 population consists of 480 markers and spans 1,386 cM (Bednarek et al. 2003). A rye consensus map combining 12 individual maps has been produced and this map contains a total of 413 markers (Börner and Korzun 1998).

7.2.5 Triticale Genetic Linkage Maps

González et al. (2005) constructed a genetic linkage map of 356 markers from a cross between the triticale varieties Torote and Presto, using RAPD, AFLP, random amplified microsatellite polymorphic (RAMP), and SSR markers in order to identify QTL for androgenetic response. The map covered a length of 2,465 cM. A high level of polymorphism was found within the R genome – the authors point out that this is consistent with the allogamous nature of rye, the donor of this genome.

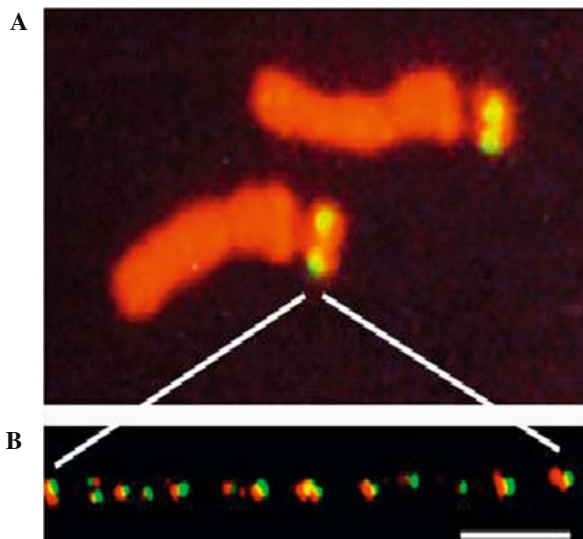
7.3 Physical Linkage Maps

A number of technologies have been developed to complement the genetic linkage maps of the Triticeae. Radiation induced deletion mutations can be used to fine-map a region of the genome to assign markers of interest to a particular phenotype. A recent example by Al-Kaff et al. (2008) localized a *cdk*-like gene cluster to the region of chromosome 5B (Ph1) controlling meiotic chromosome pairing. Radiation induced fragmentation of DNA has also been used in radiation hybrid mapping. This approach was first developed in animal cells and has now been applied to wheat (Hossain et al. 2004; Kalavacharia et al. 2006) and barley (Wardrop et al. 2002, 2004). In the case of barley, the radiation hybrid panels have been generated in a way that is similar to that in the animal area, namely, by using cell culture and fusion. X-ray irradiated barley cells (carrying the *bar* selectable marker) have been fused with tobacco cells in order to select hybrids carrying the barley genome segments generated by the high dosage of X-rays (Wardrop et al. 2002). DNA isolated from hybrid cells generate a DNA panel that is then assayed for the presence/absence of markers and markers that are physically close to each other tend to occur together more often than ones that are far apart. In this way a map order is generated with resolution depending on the number of breaks that have occurred in the DNA as a result of the X-ray irradiation. In wheat the availability of genetic stocks monosomic for distinct chromosome have been used (Hossain et al. 2004 and Kalavacharia et al. 2006). In this case an aneuploid alloplasmic durum wheat line carrying essentially an entire 1D chromosome was X-ray irradiated and these plants were crossed with a normal durum line (AABB). DNA from the progeny lines provide the radiation hybrid panels that are then assayed for markers that distinguish the D genome chromosome segments from the A and B genome chromosomes. DNA markers more often transmitted together are considered to be physically closer together than markers that show independent segregation. Radiation hybrid DNA panels have now been generated using this system for several chromosomes.

Different approaches to physical mapping have involved studying the DNA fiber. In rice, optical mapping (Zhou et al. 2007) has been developed but has not yet been applied to members of the Triticeae but appears to be of potential interest. The optical mapping of DNA involves the immobilization of isolated DNA on coated glass slides and digesting with restriction endonucleases after fixation with paraformaldehyde (Yokota et al. 1997). Visualization of the stained DNA molecules reveals gaps where digestion occurred and allows the spacing of restriction endonuclease sites to be determined. The optical mapping of the rice genome (Zhou et al. 2007) used high molecular weight DNA spread on glass slides covered with a thin layer of acrylamide and digestion with the restriction endonuclease *Swa*I. The mapping of the genome-wide distribution of restriction endonuclease sites validated the genome sequence assembly of rice. The technique of fiber Fluorescent in situ hybridization (FISH) also examines the actual DNA molecules and uses specific probes to hybridize to

high molecular weight DNA deposited on a glass slide (Fransz et al. 1996) in order to provide a physical order of sequences in the genomic DNA. The technique has been applied to the Triticeae and Fig. 7.1 shows an example from Yamamoto and Mukai (2005) where the approach was used to define the *Sec1* locus and count the number of genes located at a single site in the genome.

Fig. 7.1 Physical mapping of the *Sec-1* locus of rye. (A) Standard *in situ* hybridization of a fluorescent probe (yellow) from the *Sec-1* locus to mitotic chromosome of rye. DNA is stained red. (B) Fiber-FISH to DNA spread on a glass slides (Yamamoto and Mukai 2005). The *Sec-1* locus coding regions were detected by the green signal and the red signals were from a probe produced from the 1 kb sequence immediately upstream from the gene. Reproduced with permission from Professor Y Mukai (See Color Insert)



HAPPY mapping is an *in vitro* technique for determining the linkage between molecular markers assayed using PCR. Genomic DNA to be analyzed is diluted so that a panel of aliquots is generated that contain less than a “genome’s worth” of DNA (Thangavelu et al. 2003) in preparation for a PCR based analysis of markers of interest. Linkage between markers is detected through co-location in a given aliquot of the panel. For genomes such as wheat and barley where the physical distance between markers can be large (>50 kb), the DNA fragments need to be sufficiently long to assay linkage. It is estimated that the analysis can detect linkage between markers that are separated by approximately 0.8 times the mean length of the DNA fragments.

For a more detailed discussion of physical mapping see Chapter 11(Stein).

7.4 Map Curation

The quality of genetic maps can be affected by a number of factors, including the choice of parents, the type and the size of the population, population development and maintenance, the identification of lines and samples, sampling of plant tissue, DNA extractions, assays of marker genotypes and recording

and processing of data. The number of markers that can be placed with great accuracy onto linkage maps depends to a large extent on the size of the mapping population. Smaller mapping populations provide less chance for recombination between two markers and therefore increase the statistical uncertainty associated with ordering markers that are close to each other (Wu et al. 2003). There is no perfect algorithm or analysis method for ordering loci on linkage maps. Different methods may provide different orders, and most methods may provide multiple alternative orders in repeated analyses. Therefore, there is always some level of uncertainty associated with any marker order, especially when markers occur in clusters, i.e. large number of markers located within 10 cM distance or less. Ordering of loci typically is a multi-step, iterative process. An initial ordering step may be followed by repeated rounds of data verification and adjustment of locus order – the iterative process generally provides the basis for identifying unstable regions of the map where more work is required to more accurately define an order for the markers (Mester et al. 2003a). The final step in the construction of a map is the assessment of the genotypic data within a linkage group (the ‘graphical genotype’) to determine the quality of the map. Genetic maps are usually verified by comparisons to reference maps, including consensus maps.

Map curation ensures that map quality issues are identified and, where possible, resolved. The inspection of marker genotype data after map construction can reveal a number of issues including: (1) Progeny lines that have identical marker genotypes. As it is unlikely for two random lines to be genetically identical at a large number of loci, such lines may have been inadvertently replicated during the DH generation or mapping process; (2) Inclusion of selfed parental lines in the progeny; (3) Markers that have been scored separately but, at the level of resolution afforded by the population size, represent the same locus (e.g. markers with identical segregation data); (4) Errors in data recording, including “frame shifts” in the progeny order, particularly for marker data that are recorded manually (example illustrated in Fig. 7.2); (5) Allele phase errors, i.e. reversal of parental genotype codes (example illustrated in Fig. 7.2); (6) Mixed lines, e.g. more than one allelic pattern is observed in lines that should be homozygous; (7) Segregation distortion, which may be due to real biological causes or errors/biases in marker assays. Chromosome regions with all markers showing segregation distortion point to biological causes (e.g. due to alien introgression or selective advantage). Individual distorted markers in otherwise non-distorted regions need to be checked for scoring bias. Individual markers with strong segregation distortion may not show linkage with other markers or may appear to be linked with other similarly distorted markers not necessarily from the same linkage group; (8) Systematic patterns of crossovers such as clustering of crossovers, and markers and/or lines with high frequencies of double crossovers may indicate genotyping errors. Double crossovers, also termed ‘singletons’, occur where one line shows an allele from one parent flanked on both sides by markers showing the opposite allele; thus the locus seems to have recombined with both neighbouring loci (Lehmensiek et al. 2005

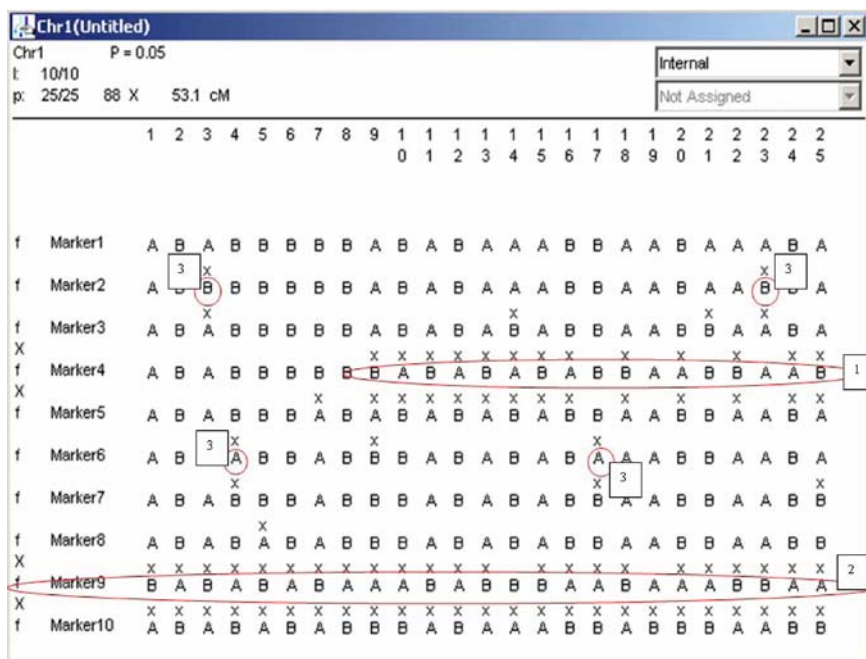


Fig. 7.2 Simulated chromosome with 10 markers showing genotyping data for 25 individuals. Examples of errors in data recording are indicated: (1) “frameshifts” (2) allele phase errors (3) double crossovers (*See Color Insert*)

and Van Os et al. 2005b). In most population types there should not be more than 2–3 recombination events per chromosome. Genotypes need to be checked for errors and individual genotype scores may need to be replaced by the missing data symbol, taking into account map distances (Hackett and Broadfoot 2003 and Lehmensiek et al. 2005). Genotyping errors lead to unexpected crossovers, which depending on the total number of markers in a map can result in large increases of map length (examples illustrated Fig. 7.2); (9) Alien segments may cause problems as very little or no recombination may occur in these segments, thus making the ordering of markers impossible. In such cases the marker order can only be estimated by comparison to consensus or other maps.

7.5 Consensus Maps

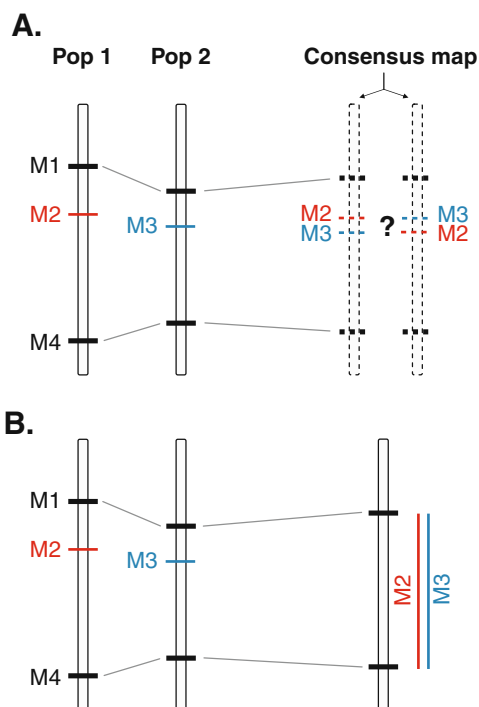
Depending on the genetic distance between parents, only a subset of the markers that are theoretically available for a given species are expected to segregate in any bi-parental cross. To place the majority of available markers on a single ‘consensus’ map it is therefore necessary to jointly analyse mapping

data from several populations. There are several programs that can be used to build consensus maps, including JoinMap, Carthagene and PhenoMap (Stam 1993; De Givry et al. 2005, GeneFlow Inc., Alexandria, VA, USA). JoinMap, however, which jointly considers all available recombination frequencies between marker pairs (averaged across populations), seems to be the most widely used program, perhaps because of its convenient graphical user interface (other programs capable of building consensus maps such as Carthagene have predominantly command line-driven interfaces). Our experience, however, suggests that the regression algorithm of JoinMap can produce erroneous marker orders when building consensus maps from high-density datasets, even if the recombination frequencies in different populations appear to be homogenous (a pre-requisite for merging data sets from different populations with the regression algorithm of JoinMap). Importantly, this problem also affects the ordering of markers in individual maps and was described in detail by Wenzl et al. (2006). The problem was not related to the quality of marker scores because misplaced markers frequently co-segregated with markers in other chromosome areas. In addition, the regression algorithm excluded some markers that perfectly fitted (i.e., co-segregated) with markers that were incorporated into the map. These problems may be the result of an inherent limitation of the regression algorithm in dealing with large numbers of closely linked markers. If the graphical genotype underlying a map is not inspected, an erroneous marker order can easily go unnoticed. We suggest that the marker scores should always be released into the public domain when a new map is published, either as a supplementary file or through deposition at a public database such as GrainGenes (<http://wheat.pw.usda.gov>). The practice of publishing raw data is well established for microarray experiments, which produce a far greater volume of data. It should become standard for genetic maps as well.

A more recent version of JoinMap (version 4) has an alternative marker-ordering algorithm to improve ordering of closely linked markers. However, this algorithm is only available for individual maps, and the regression algorithm continues to be the only option for building a consensus map. Inspecting the graphical genotypes of individual populations (with markers ordered according to the consensus order reported by JoinMap) thus remains a crucial step in the process of building a consensus map. If misplaced markers are identified, the analysis may have to be repeated after varying the program's settings and/or excluding some markers. The marker order may even have to be curated manually to some degree (Wenzl et al. 2006).

An important point to keep in mind when building a consensus map is that only a subset of markers can be ordered by directly comparing their segregation signatures (see markers M1 and M4 in Fig. 7.3A). Other markers can only be placed on the map by comparing their positions relative to the positions of 'bridging' markers that segregate in more than one population (markers M2 and M3 in Fig. 7.3A). The confidence with which markers are positioned on a consensus map, therefore, varies from marker to marker. These different degrees of uncertainty, however, are usually ignored when graphically

Fig. 7.3 Challenge of optimising the order of markers in consensus maps. (A) “Traditional” approach: recombination frequencies measured in different populations are used to establish the relative order of all markers, including those that were not genotyped in the same population. (B) Alternative, probably more realistic way to represent marker order in consensus maps which takes into account the various degrees of uncertainty associated with the position of individual markers (*See Color Insert*)



representing a consensus map, perhaps because of a lack of tools to quantify and display them. We note that the latter point has been addressed in Yap et al. (2003) but their approach has not been widely adopted to date. A simple graphical representation of a consensus map, tends to be misleading because the absence of error bars for marker locations may cause the impression of a degree of precision that is not justified considering the degree of statistical uncertainty inherent in the original data. This problem clearly limits the usefulness of consensus maps.

An alternative strategy could be to assign individual markers to map regions (vertical bars) rather than precise positions (Fig. 7.3B). Such an approach would result in an assembly of partly overlapping bars of varying size, which represent the ‘positions’ of individual markers and the degree of uncertainty associated with them. The sizes of these bars would depend on several factors, amongst them the number of populations in which the markers were assayed. As more populations are added for re-analysis, the bars (map regions) representing individual markers would progressively shrink. During this process, the relative order of an increasing number of markers would be resolved, thus reflecting the increased information content of the underlying set of segregation data.

Another approach that has been considered (C. Feuillet, P Sourdille, A Korol and R Appels, unpublished) is to utilize all available maps to produce

a framework map which carries only those markers that are in a consistent order across all maps. The framework-map approach avoids forcing a marker order on a large number of markers. Particular regions of the map can then be shown in more detail as sourced from a specific high resolution map. This approach accepts that at a high resolution level, genuine differences can exist between different wheat or barley varieties used in crossing and generating progeny for map construction. A composite or consensus map will thus show the framework map and specific regions elaborated by analyses of progeny from a cross that is identified as the source of the data (Fig. 7.4).

Good-quality maps are indispensable for QTL mapping, marker-assisted selection, comparative mapping, high-resolution mapping and map-based cloning. The major use of genetic linkage maps has, however, been to identify QTL.

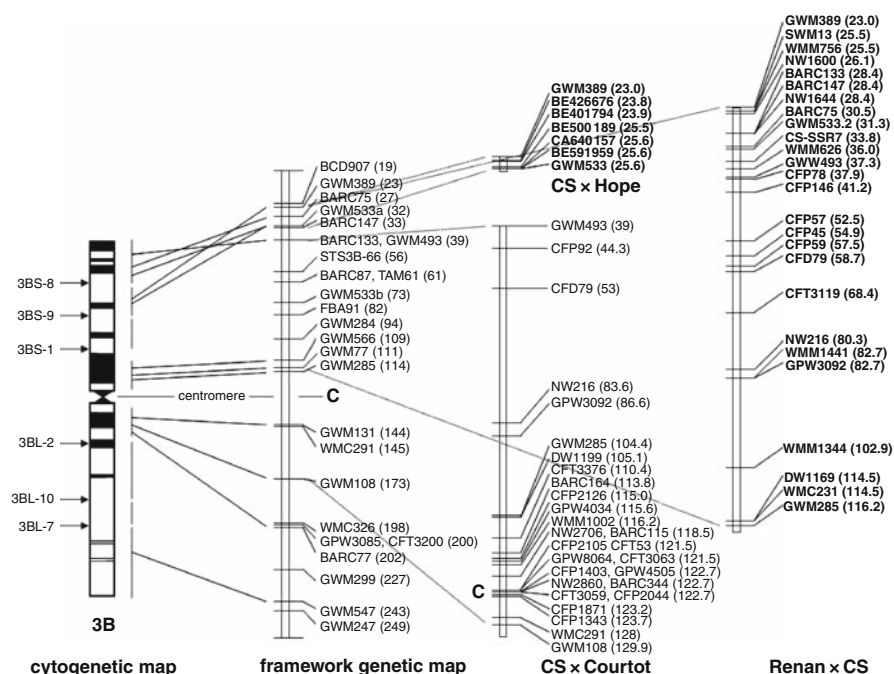


Fig. 7.4 Example of a composite or framework map for chromosome 3B of wheat. The framework genetic map was compiled from 15 published molecular genetic maps and any markers that showed an ambiguous location relative to the other markers were removed. Numbers in brackets indicated genetic distance in cM. The framework map can be aligned to the cytogenetic maps based on deletion bins defined by Endo and Gill (1996). The core set markers in this framework map are more likely to be applicable to new maps. Individual biparental crosses from large populations are identified as filling-in the detailed molecular marker content and order of specific regions of the framework genetic maps

7.6 QTL Mapping

Many traits of interest to breeding programs show a continuous range of values (for example yield and quality) rather than forming distinct phenotypic classes. Such traits are under the control of several genes (and the environment), which are referred to as polygenes or quantitative trait loci (QTL) (Tanksley 1993). Molecular markers are useful for identifying loci that control quantitative traits (Langridge et al. 2001) because markers that tend to be transmitted with specific trait values are likely to be close to a gene affecting the trait (Doerge et al. 1997).

The simplest QTL detection methods are based on ANOVA and linear regression (Hackett 2002), likelihood analyses (Doerge et al. 1997) or *t*-tests (Collard et al. 2005). For a DH population, the process involves scoring the marker data of the population, and calculating and comparing phenotypic means of the two genotypic classes to identify significant differences. If a significant difference is found, it is concluded that the marker is linked to a locus affecting the trait of interest. These types of analyses do not require a genetic linkage map, and are referred to as single point or single marker analysis. The main disadvantage of single point analysis is that the further away the marker is from the gene, the less likely it is to be detected due to recombination between the marker and the gene (Tanksley 1993).

To overcome the problems associated with single marker analysis, Lander and Botstein (1989) devised the method of interval analysis (or interval mapping). This method requires a genetic linkage map. Interval mapping builds upon single point analysis by compensating for recombination between the marker and the gene affecting the trait of interest by using genetically linked markers for the analysis (Lander and Botstein, 1989). As Tanksley (1993) points out, this method is of maximum benefit when linked markers are fairly far apart (because of a large number of recombination events), but when markers are more dense, the single point analysis gives similar results to interval mapping.

Composite interval mapping (Zeng 1994) and multiple QTL mapping (Jansen 1993) are more refined methods that may be used for QTL detection. Interval mapping assesses the likelihood of a single QTL at each location on the genome – however QTL located elsewhere on the genome can have an interfering effect (Jansen 1993). By combining interval mapping with multiple regression (using markers associated with other QTL as cofactors), these techniques are superior to interval mapping and increase the accuracy and precision of QTL mapping (Hackett 2002). Other techniques such as Bayesian interval mapping and outlier detection methods have been described recently (Yandell et al. 2007 and Verbyla et al. 2007).

7.6.1 Practical Considerations for QTL Mapping

Each of the QTL mapping methods (single marker, interval mapping, composite interval and multiple QTL mapping) can be performed using a range of

computer programs – however as Asins (2002) aptly concludes: “QTL mapping is much more than running a programme”. Asins (2002) points out other factors that should be considered, including: population size; the heritability of the trait; the number of QTL; their interaction; and the reliability of the marker order of the linkage map.

The size of the population and the heritability of the trait (the proportion of phenotypic variance that is genetic; Lynch and Walsh 1998), are the most important factors that affect QTL mapping studies (Collard et al. 2005). Tanksley (1993) suggests that with typical population sizes used for QTL mapping (100–250 individuals), only QTL with large effect are likely to be identified, and furthermore, the effect of QTL that are identified can be over-inflated. Studies using both simulated and experimental data have confirmed these hypotheses. In a simulation study, Beavis (1994) reported that phenotypic variances associated with QTL are greatly overestimated, and that this is most pronounced if only small populations (e.g. 100 individuals) are evaluated. Beavis (1994) suggested that the actual phenotypic variance explained by QTL can only be accurately estimated from populations with 500 to 1,000 individuals. Furthermore, if 10 loci affect a trait of 30% heritability and a population of 100 F_2 is used for mapping, each true QTL has only a 9% probability of being detected, and the variance explained will be overestimated by a factor of 5.6 (Beavis 1994, 1998). Melchinger et al. (1998) confirmed the low power of QTL detection and large bias of QTL effects by comparing population sizes of $N = 344$ and $N = 107$ for detection of QTL controlling various agronomic traits in maize. With the larger population, 107 QTL were detected. With the smaller population, 39 QTL were detected. Only 20 QTL were in common between the different sized populations. The authors concluded that QTL effects need to be estimated in an independent population before they can be used with any reliability in marker-assisted selection schemes. It should be noted that Lande and Thompson (1990) also suggested that a way to obtain unbiased estimates of QTL effects was to map QTL in one cross and then confirm the effects of the detected QTL in another population (so called “marker validation”). This approach not only addresses the issue of insufficient population size, but also provides an indication of the transferability of a particular QTL into an independent genetic background (Castro et al. 2003). It is likely that the recommendations of Lande and Thompson (1990) were infrequently followed because of the cost of genotyping in the past. More recently however, the cost of genotyping has decreased substantially, and validation studies are becoming a more frequent component of QTL studies.

It is clear that the heritability of the trait and the size of the mapping population will have an effect on the number of QTL that can be detected. Kearsey and Farquhar (1998) point out that, because only QTL with significant effects are reported in the literature, the phenotypic effects of the reported QTL will be biased towards larger values. These biases are larger for QTL with small effects, and thus imply that studies will tend to underestimate the true number of QTL, but exaggerate their effect (Kearsey and Farquhar 1998).

Carlborg and Haley (2004) suggest that interactions between QTL (epitasis) are often neglected in studies of complex traits. The lack of studies demonstrating epistatic interactions between QTL has been attributed to: the low statistical power in small population sizes that are typically used in QTL mapping (Tanksley 1993); the absence of biological evidence to confirm such interactions (Carlborg and Haley 2004) and; the lack of availability of appropriate statistical tests to assess their significance (Cordell 2002). However, with “whole-genome” molecular maps becoming more common (Table 7.1), investigations of the importance of epistasis are now becoming more common. In rice, for example, epistatic interactions have been identified in a range of phenotypes such as panicle number (Liao et al. 2001), yield components (Xing et al. 2002 and Zhuang et al. 2002), and heterosis (Yu et al. 1997). In wheat, epistatic interactions appear important for the effects of glutenin loci on dough rheological properties (Ma et al. 2005). Each of these studies has found that the contribution to phenotypic variance of the main-effect QTL is larger than that of epistatic QTL – nonetheless, the apparent ubiquity of epistatic interactions warrants further investigation of this phenomenon in QTL mapping studies.

The ordering of loci within linkage groups is of great importance for robust QTL detection. Wu et al. (2003) have suggested that even if QTL detection methods are appropriate, QTL identified may be incorrect if marker order is inaccurate. Indeed, in a recent study investigating the effect of thorough map curation, Lehmensiek et al. (2005) found that reordering of marker loci not only improved QTL resolution, but also affected the magnitude of QTL effects. In contrast to the findings of Wu et al. (2003) and Lehmensiek et al. (2005), Dodds et al. (2004) found in simulation studies that the accuracy of the map had little or no impact on the detection of QTL, provided no markers are assigned to an incorrect linkage group. The findings of Dodds et al. (2004) are surprising given the number of reports that emphasize the importance of an accurate marker order for QTL mapping (Asins 2002; Lehmensiek et al. 2005; Van Os et al. 2005a,b; Wu et al. 2003). The reason for this apparent discrepancy may be that Dodds et al. (2004) only investigated the effects of one or two markers being ordered incorrectly.

7.7 High-Resolution Mapping

Quantitative genes are usually identified on a 10–30 centimorgan segment of a chromosome, thus making it virtually impossible to pinpoint and clone the gene directly. High-resolution mapping is subsequently used to reduce the confidence interval for a QTL and to obtain markers closer to the gene. Populations of 1,000 or more individuals are used for this purpose, in order to provide sufficient genetic resolution in the gene region. The degree of polymorphism and the recombination frequency within the chromosomal segment under study are dependent on the parents chosen for the population. The parental lines are therefore crucial, and in some cases more than one population is used (Pellio et al. 2005).

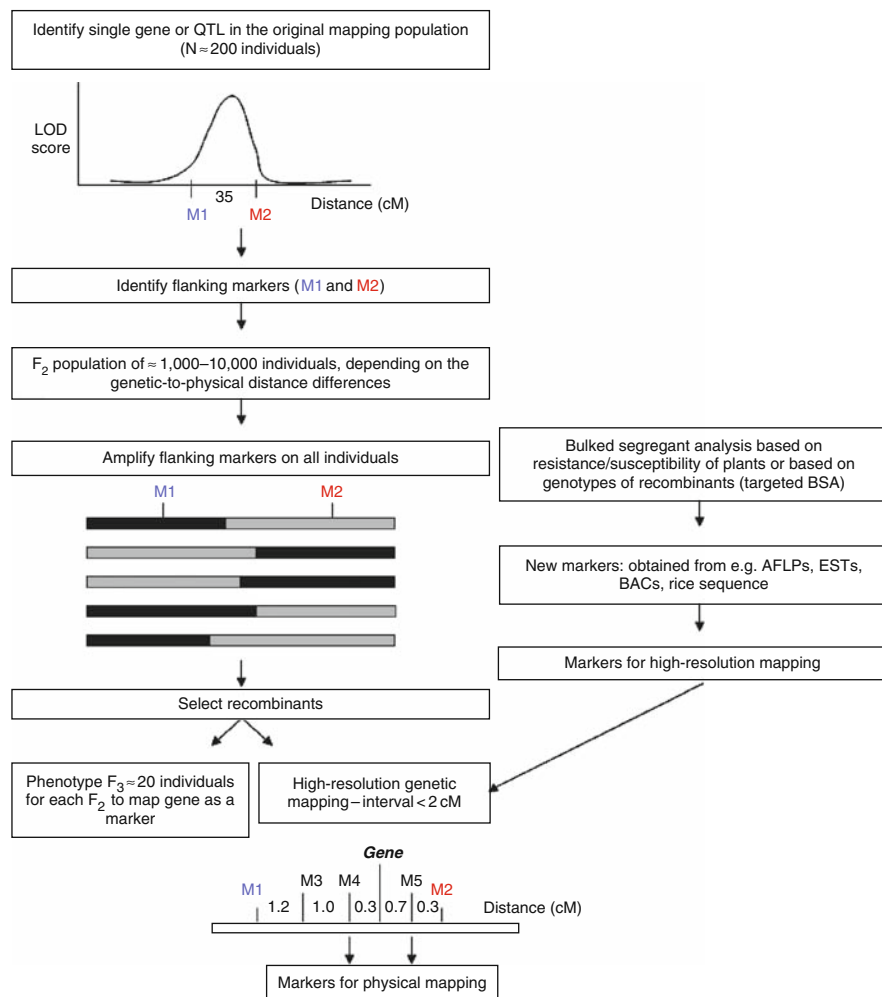


Fig. 7.5 This figure illustrates the key steps involved in high-resolution mapping. Initially a gene or QTL is localized in a mapping population of approximately 200 segregating lines. Markers flanking the gene are identified and a F₂ population of between 1,000 and 10,000 segregating lines is produced for fine mapping. The two flanking markers are amplified on all lines of the population and recombinant lines selected. Phenotypes of the selected recombinant lines are determined by phenotyping about 20 F₃ lines for each F₂ individual. The flanking marker interval is screened for polymorphic markers. These markers are mapped on the selected recombinant lines with the aim of reducing the genetic map interval containing the gene to less than 1 cM. The gene may then be close enough to be selected from a BAC clone

The key steps involved in high-resolution mapping are illustrated in Fig. 7.5. After the initial localization of a gene or QTL in a mapping population of approximately 200 segregating lines, markers flanking the gene are identified. The interval between the flanking markers typically ranges between 2 and

Table 7.2 Summary of some of the high-resolution mapping studies conducted in the Triticeae. The species studied and the trait/gene investigated are given together with the number of plants screened and the number of recombinants identified

Species	Trait	Name of Gene	F ₂ plants screened	Recombinants	Reference
Wheat	Protein content	<i>GPC-B1</i>	935	291	Distelfeld et al. 2006
Wheat	Sensitivity toxin (Ptr ToxA) produced by tan spot fungus	<i>Tsn1</i>	6500		Haen et al. 2004
wheat	Sensitivity toxin (Ptr ToxA) produced by tan spot fungus	<i>Tsn1</i>	2719		Lu et al. 2006
wheat	Leaf rust resistance	<i>Lr1</i>	3658		Ling et al. 2003
wheat	Leaf rust resistance	<i>Lr10</i>	3120	96	Stein et al. 2000
barley	Aluminium tolerance	<i>Alp</i>	2070	236	Wang et al. 2007
barley	Boron-toxicity tolerance	<i>Bot1</i>	3360		Sutton et al. 2007
barley	Stem rust resistance	<i>Rpg1</i>	4259	17	Brueggeman et al. 2002
barley	Leaf rust resistance	<i>Rdg2a</i>	1400	93	Bulgarelli et al. 2004
barley	Six-rowed spike	<i>vrs1</i>	2212	6	He et al. 2004
barley	Naked caryopsis	<i>nud</i>	869	1	Taketa et al. 2006
barley	Barley yellow mosaic virus	<i>Rym4, Rym5</i>	4924	144	Pellio et al. 2005
rye	Aluminum tolerance	<i>Alt3</i>	1123	15	Miftahudin et al. 2005
rye	Rust resistance	<i>Sr31, Lr26, Yr9</i>	1580	36	Mago et al. 2005

30 cM. F₂ populations of between 1,000 and 10,000 segregating lines are usually produced for fine mapping (Table 7.2). The two flanking markers are genotyped on all lines of the population, and lines recombinant for the region between the flanking markers are selected. Phenotypes of the selected recombinant lines are determined by phenotyping about 20 F₃ lines for each F₂ individual. Bulk segregant analysis (BSA, Michelmore et al. 1991) is used to screen the flanking marker interval for polymorphic markers. Plants used in the bulks are selected based on either the absence or presence of the expression of the trait or the genotypes of recombinants (targeted BSA). Many molecular markers may have to be screened to obtain markers that are polymorphic in the target region. Amplified fragment length polymorphism (AFLP) markers have mainly been used for this purpose (Faris and Gill 2002; Gu et al. 2004; Haen et al. 2004; Mago et al. 2005; Peleman et al. 2005; Pellio et al. 2005; Yang et al. 2004).

However, EST-based markers have been used to increase the marker density around the *Tsn1* locus in wheat (Lu et al. 2006). New markers in the target region have also been detected by sequencing fragments of bacterial artificial chromosomes (BACs) or by making use of the rice sequence (Brueggeman et al. 2002; Miftahudin et al. 2005; Ling et al. 2003; Stein et al. 2000; Sutton et al. 2007; Wang et al. 2007). A recently developed technique, High-resolution melting analysis (HRM), which characterizes nucleic acid samples based on their dissociation or melting behaviour may also be a useful approach in this context. HRM can detect differences between samples according to their sequence (including SNPs), length, GC content and strand complementarity (Liew et al. 2004; Montgomery et al. 2007; Lehmensiek et al. 2008).

Newly identified markers are mapped on the selected recombinant lines with the aim of reducing the genetic map interval containing the gene to less than 1 cM. Considering that 100 kb is the average insert size of a BAC clone, ideally a resolution of one recombination at every 100 kb is needed to have flanking markers close enough to the gene to be able to select the gene from a BAC clone (Pellio et al. 2005). Therefore, assuming evenly distributed recombination events, one needs approximately 5,500, 4,000 or 287 meiotic events to achieve 100-kb resolution in wheat, barley and rice, respectively. As recombination in most species is not evenly distributed and mainly confined to a few relatively small areas spaced by large segments with severely suppressed recombination, such as the centromeric regions (Chen et al. 2002; Faris et al. 2000; Gill et al. 1996; Künzel et al. 2000; Stein et al. 2000), the number of meiotic events needed for high-resolution mapping is dependent on the chromosomal region under study and may vary quite drastically. For example, a detailed genetic and physical map constructed for a highly recombinant region on chromosome 5BL in wheat indicated that this region accounted for 4% of the physical size of the long arm and at least 30% of the recombination along the whole chromosome (Faris et al. 2000). Multiple cross-overs occurred in this region and the recombination in this region was at least 11 times greater than the genomic average. Similarly, in barley the regions of highest recombination frequency, i.e. ≤ 1 Mb/cM correspond to only 4.9% of the total genome and recombination in these regions is 20 times that of the remaining genome (Künzel et al. 2000). In wheat, Faris et al. (2003) observed physical-to-genetic distance differences ranging from 130 kb/cM to 600 kb/cM on a contig spanning 0.9 cM. For high-resolution mapping of the *Rym4/Rym5* locus in the telomeric region of chromosome arm 3HL in barley only 500–1,000 meiotic events would have been enough to obtain the genetic resolution needed as a ratio of 400 kb/cM was determined for this region (Pellio et al. 2005). More than 3,000 F_2 lines were used to construct a BAC contig spanning the *Lr10* rust resistance locus in wheat (Stein et al. 2000). On the other hand, an F_2 population of only 465 lines was used to construct a BAC contig spanning the *Q* gene region in wheat (Faris et al. 2003).

In plant studies, F_2 populations are favoured for high-resolution mapping as they are easy and fast to generate. Populations such as nearly isogenic lines

(NILs) that differ in genetic composition only at the segment harbouring the QTL can also be used, but this approach requires many generations of crosses. QTL or genes can also be fine mapped by association mapping using large numbers of related pedigrees. Nullisomic-tetrasomic and deletion lines have been used to find markers closer to the Q gene on chromosome arm 5AL in wheat (Faris and Gill 2002). A F₂ population consisting of 190 lines of a cross between two Chinese Spring varieties, one of them containing a pair of *T. dicoccoides* 5A chromosomes substituted for the native 5A chromosomes, was used for genetic mapping in this case. High resolution mapping is an essential component for any positional cloning effort because it provides researchers with flanking markers that are closer to the gene of interest.

7.8 Future Directions

The production of genetic maps of wheat, barley and other Triticeae species is moving towards higher resolution mapping of specific regions associated with important agronomic or quality traits. Coincident with the need for higher resolution is a trend toward analyzing large F₂ populations because of the high cost of producing DH lines. The new marker technologies discussed in this chapter and the chapter by Paux and Sourdille (Chapter 9) are a significant factor in facilitating the analysis of these large populations. The primary purpose of high-resolution mapping is to align a physical/sequence map with the respective region of the genetic map in order to identify new probes and candidate genes that can be used to investigate fundamental features of chromosome regions encoding the traits of interest.

Novel approaches such as pedigree-based mapping and the use of integrated multiple crosses (MAGIC) are likely to gradually replace traditional approaches based on bi-parental crosses because they expand the 'genotype space' that can be explored. The MAGIC approach is particularly communal in that it maximizes the number of phenotypes that are segregating in a single large (approximately 5,000) RIL population. MAGIC populations should not only be useful for mapping target loci but are also an excellent tool for high-resolution mapping and cloning of trait-related genes.

To date, 1,051,768 wheat and 478,734 barley ESTs (June 2008; http://www.ncbi.nlm.nih.gov/dbEST/dbEST_summary.html) have been sequenced and EST collections are expanding rapidly to provide new opportunities for SNP-based markers. These in turn provide improved genome coverage for molecular genetic maps. Currently a number of SNP maps are being developed for barley (<http://harvest.ucr.edu>) and wheat (<http://rye.pw.usda.gov/snpworld/Map>).

Genetic maps are the fundamental tool to identify features of phenotypes that are linked to specific genetic loci and eventually DNA sequences or genes. The broad range of available software options for building maps and

identifying QTL provide a good foundation for further advances in genetic analysis. As molecular marker and phenotyping technologies evolve towards high-throughput platforms, the increasing number and size of datasets call for a higher degree of automation and standardization in several critical areas. These areas include sample handling procedures (seeds, DNA), data exchange and storage, genetic-map construction and curation, QTL analysis and trait ontologies. Map construction and QTL analysis are usually iterative processes where the output of one step serves as an input for a refined analysis (e.g. after excluding potential errors or focusing on a subset of data). These iterations are typically performed manually by users, but could be encoded in more advanced genetic-analysis software. Given the availability of a broad spectrum of alternative QTL-mapping methods, a 'QTL pipeline' could encompass a suite of complementary algorithms and include a built-in validation of algorithm performances using simulated datasets that resemble the QTL models derived from the data (Guoyou Ye, personal communication). However, it is equally important to build and reinforce the capacity for understanding and analyzing genetic data at a time when the generation of raw data (genotyping, phenotyping, sequence analyses) is increasingly out-sourced.

The ultimate aim for many aspects of cell biology is the characterization of genetic networks that account for variation in a particular phenotype, and genetic maps are clearly a vital tool to bridge the gap between phenotype and the genetic makeup of plants.

References

- Aishwarya, V. and Sharma, P.C. (2007) UgMicroSatdb: database for mining microsatellites from unigenes. *Nucl Acids Res.* doi:10.1093/nar/gkm811, 1–4.
- Akbari, M., Wenzl, P., Caig, V., Carling, J., Xia, L., Yang Suszynski, G., Mohler, V., Lehmensiek, A., Kuchel, H., Hayden, M., Howes, N., Sharp, P., Vaughan, P., Rathmell, W., Huttner, E. and Killian, A. (2006) Diversity arrays technology (DArT) for high throughput profiling of the hexaploid wheat genome. *Theoretical and Applied Genetics* 113, 1409–1420.
- Al-Kaff, N., Knight, E., Bertin, I., Foote, T., Hart, N., Griffiths, S. and Moore, G. (2008) Detailed dissection of the chromosomal region containing the Ph1 locus in wheat *Triticum aestivum*; with deletion mutants and expression profiling. *Annals of Botany* 101, 863–872.
- Anderson, O.D., Greene, F.C., Yip, R.E., Halford, N.G., Shewry, P.R. and Malpica-Romero, J.M. (1989) Nucleotide sequences of the two high-molecular-weight glutenin genes from the D genome of hexaploid wheat, *Triticum aestivum* L. cv. Cheyenne. *Nucleic Acids Research* 17, 461–462.
- Appels, R. (2003) A consensus molecular genetic map for wheat – a cooperative international effort. In: Prognia, N.E. (ed.), *Proceedings of the 10th International Wheat Genetics Symposium*, 211–214.
- Appels, R., Francki, M. and Chibbar, R. (2003) Advances in cereal functional genomics. *Functional and Integrative Genomics* 3, 1–24.
- Asins, M.J. (2002) Present and future of quantitative trait locus analysis in plant breeding. *Plant Breeding* 121, 281–291.

- Badr, A., Miller, K.M., Schafer-Preg, R., El Rabey, H., Effgen, S., Ibrahim, H.H., Pozzi, C., Rhode, W. and Salamini, F. (2000) On the origin and domestication history of barley (*Hordeum vulgare*). *Molecular Biology and Evolution* 17, 499–510.
- Beavis, W.D. (1994) The power and deceit of QTL experiments: lessons from comparative QTL studies. Paper presented at the Proceedings of the Forty-ninth Annual Corn and Sorghum Industry Research Conference, Washington, D.C.
- Beavis, W.D. (1998) QTL analyses: power, precision, and accuracy. In: Paterson, A.H. (ed.) *Molecular Dissection of Complex Traits*. CRC Press, Boca Raton.
- Bednarek, P.T., Masojø, P., Lewandowska, R. and Myeoków, B. (2003) Saturating rye genetic map with amplified fragment length polymorphism (AFLP) and random amplified polymorphic DNA (RAPD) markers. *Journal of Applied Genetics* 44, 21–33.
- Beecher, A., Bettge, A., Smidansky, E. and Grioux, M.J. (2002) Expression of wild-type pinB sequence in transgenic wheat complements a hard phenotype. *Theoretical and Applied Genetics* 105, 870–877.
- Bonafede, M., Kong, L., Tranquilli, G., Ohm, H. and Dubcovsky, J. (2006) Reduction of a *Triticum monococcum* chromosomes segment carrying the softness genes *Pina* and *Pinb* translocated to bread wheat. *Crop Science* 47, 821–828.
- Börner, A. and Korzun, V. (1998) A consensus linkage map of rye (*Secale cereale* L.) including 374 RFLPs, 24 isozymes and 15 gene loci. *Theoretical and Applied Genetics* 97, 1279–1288.
- Brueggeman, R., Rostoks, N., Kudrna, D., Kilian, A., Han, F., Chen, J., Druka, A., Steffenson B. and Kleinhofs, A. (2002) The barley stem rust-resistance gene *Rpg1* is a novel disease-resistance gene with homology to receptor kinases. *Proceedings of the National Academy of Sciences of the United States of America* 99, 9328–9333.
- Bulgarelli, D., Collins, N.C., Tacconi, G., Dellaglio, E., Brueggeman, R., Kleinhofs, A., Stanca, A.M. and Vale, G. (2004) High-resolution genetic mapping of the leaf stripe resistance gene *Rdg2a* in barley. *Theoretical and Applied Genetics* 108, 1401–1408.
- Carlborg, O. and Haley, C.S. (2004) Epistasis: too often neglected in complex trait studies? *Nature Reviews Genetics* 5, 618–625.
- Castiglioni, P., Pozzi, C., Heun, M., Terzi, V., Muller, K.J., Rohde, W. and Salamini, F. (1998) An AFLP-based procedure for the efficient mapping of mutations and DNA probes in barley. *Genetics* 149, 2039–2056.
- Castro, A.J., Chen, X., Corey, A., Filichkina, T., Hayes, P., Mundt, C., Richardson, K., Sandoval-Islas, S. and Vivar, H. (2003) Pyramiding and validation of quantitative trait locus (QTL) alleles determining resistance to barley strip rust: Effects of Adult Plant Resistance. *Crop Science* 43, 2234–2239.
- Cavannah, C., Morell, M., Mackay, I. and Powell, W. (2008) From mutations to MAGIC: resources for gene discovery, validation and delivery in crop plants. *Current Opinion in Plant Biology* 11, 1–7.
- Chalmers, K.J., Campbell, A.W., Kretschmer, J., Karakousis, A., Henschke, P.H., Pierens, S., Pallotta, M., Cornish, G.B., Shariflou, M.R., Rampling, L.R., McLauchlan, A., Daggard, G., Sharp, P.J., Holton, T.A., Sutherland, M.W., Appels, R. and Langridge, P. (2001) Construction of three linkage maps in bread wheat, *Triticum aestivum*. *Australian Journal of Agricultural Research* 52, 1089–1119.
- Chao, S., Sharp, P.J., Worland, A.J., Warham, E.J., Koebner, R.M.D. and Gale, M.D. (1989) RFLP-based genetic maps of wheat homoeologous group 7 chromosomes. *Theoretical and Applied Genetics* 78, 495–504.
- Chartrain, L., Brading, P.A., Widdowson, J.P. and Brown, J.K.M. (2004) Partial resistance to *Septoria tritici blotch* (*Mycosphaerella graminicola*) in wheat cultivars Arina and Riband. *Phytopathology* 94, 497–504.

- Chen, J., Griffey, C.A., Saghai Maroof, M.A., Stromberg, E.L., Biyashev, R.M., Zhao, W., Chappell, M.R., Pridgen, T.H., Dong, Y. and Zeng, Z. (2006) Validation of two major quantitative trait loci for fusarium head blight resistance in Chinese wheat line W14. *Plant Breeding* 125, 99–101.
- Chen, M., Presting, G., et al. (2002) An integrated physical and genetic map of the rice genome. *The Plant Cell* 14, 537–545.
- Collard, B.C.Y., Jahufer, M.Z.Z., Brouwer, J.B. and Pang, E.C.K. (2005) An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. *Euphytica* 142, 169–196.
- Cordell, J. (2002) Epistasis: what it means, what it doesn't mean, and statistical methods to detect it in humans. *Human Molecular Genetics* 11, 2463–2468.
- Costa, J.M., Corey, A., Hayes, P.M., Jobet, C., Kleinhofs, A., Kopisch-Obusch, A., Kramer, S.F., Kudrna, D., Li, M., Riera-Lizarazu, O., Sato, K., Szucs, P., Toojinda, T., Vales, M.I. and Wolfe, R.I. (2001) Molecular mapping of the Oregon Wolfe Barleys: a phenotypically polymorphic doubled-haploid population. *Theoretical and Applied Genetics* 103, 415–424.
- De Bustos, A. and Jouve, N. (2002) Phylogenetic relationships of the genus *Secale* based on the characterisation of the rDNA ITS sequences. *Plant Systematics and Evolution* 235, 147–154.
- De Givry, S., Bouchez, M., Chabrier, P., Milan, D. and Schiex, T. (2005) CARTHAGENE: multipopulation integrated genetic and radiated hybrid mapping. *Bioinformatics* 21, 1703–1704.
- Devos, K.M., Atkinson, M.D., Chinoy, C.N., Liu, C.J. and Gale, M.D. (1992) RFLP-based genetic map of the homoeologous group 3 chromosomes of wheat and rye. *Theoretical and Applied Genetics* 83, 931–939.
- Devos, K.M. and Gale, M.D. (1992) The use of random amplified polymorphic DNA markers in wheat. *Theoretical and Applied Genetics* 84, 567–572.
- Devos, K.M., Atkinson, M.D., Chinoy, C.N., Francis, H.A., Harcourt, R.L., Koebner, R.M.D., Liu, C.J., Masojc, P., Xie, D.X. and Gale, M.D. (1993) Chromosomal Rearrangements in the Rye Genome Relative to That of Wheat. *Theoretical and Applied Genetics* 85, 673–680.
- Distelfeld, A., Uauy, C., Fahima, T., Dubcovsky, J. (2006) Physical map of the wheat high-grain protein content gene *Gpc-B1* and development of a high-throughput molecular marker. *New Phytologist* 169, 753–763.
- Dodds, K.G., Ball, R., Djorovic, N. and Carson, S.D. (2004) The effect of an imprecise map on interval mapping QTLs. *Genetic Research Cambridge* 84, 47–55.
- Doerge, R.W., Zeng, Z.-B. and Weir, B.S. (1997) Statistical issues in the search for genes affecting quantitative traits in experimental populations. *Statistical Science* 12, 195–219.
- Dubcovsky, J. (2004) Marker-assisted selection in public breeding programs: The wheat experience. *Crop Science* 44, 1895–1898.
- Endo, T.R. and Gill, B.S. (1996) The deletion stocks of common wheat. *Journal of Heredity* 87, 295–307.
- Erayman, M., Sandhu, D., Sidhu, D., Dilbirligi, M., Baenziger, P.S. and Gill, K.S. (2004) Demarcating the gene-rich regions of the wheat genome. *Nucleic Acids Research* 32, 3546–3565.
- Faris, J.D., Fellers, J.P., Brooks, S.A. and Gill, B.S. (2003) A bacterial artificial chromosome contig spanning the major domestication locus Q in wheat and identification of a candidate gene. *Genetics* 164 311–321.
- Faris, J.D. and Gill, B.S. (2002) Genomic targeting and high-resolution mapping of the domestication gene *Q* in wheat. *Genome* 45, 706–718.
- Faris, J.D., Haen, K.M. and Gill, B.S. (2000) Saturation mapping of a gene-rich recombination hot spot region in wheat. *Genetics* 154, 823–835.

- Francki, M. and Appels, R. (2007) Comparative genomics and crop improvement. In: Brown, J.R. (ed.), *Comparative genomics: basic and applied research*. CRC press, Boca Raton, London, New York, pp 321–340.
- Fransz, J.P. C., Liharska, J.B., Peeters, A.J., Zabel, P. and de Jong, J.H. (1996) High-resolution physical mapping in *Arabidopsis thaliana* and tomato by fluorescence in situ hybridization to extended DNA fibres. *The Plant Journal* 9, 421–430.
- Gill, K.S., Gill, B.S., Endo, T.R. and Mukai, Y. (1993) Fine physical mapping of Ph1, a chromosome pairing regulator gene in polyploid wheat. *Genetics* 134, 1231–1236.
- Gill, K.S., Gill, B.S., Endo, T.R. and Boyko, E.V. (1996) Identification and high-density mapping of gene-rich regions in chromosome group 5 of wheat. *Genetics* 143, 1001–1012.
- González, J.M., Muniz, L.M. and Jouve, N. (2005) Mapping of QTLs for androgenetic response based on a molecular genetic map of \times Triticosecale Wittmack. *Genome* 48, 999–1009.
- Graner, A., Jahoor, A., Schondelmaier, J., Siedler, H., Pillen, K., Fischbeck, G., Wenzel, G. and Herrmann, R.G. (1991) Construction of an RFLP map of barley. *Theoretical and Applied Genetics* 83, 250–256.
- Gu, K., Tian, D., Yang, F., Wu, L., Sreekala, C., Wang, D., Wang, G.L. and Yin, Z. (2004) High-resolution genetic mapping of *Xa27(t)*, a new bacterial blight resistance gene in rice, *Oryza sativa* L. *Theoretical and Applied Genetics* 108, 800–807.
- Gupta, P.K., Roy, J.K. and Prasad, M. (2001) Single nucleotide polymorphisms: A new paradigm for molecular marker technology and DNA polymorphism detection with emphasis on their use in plants. *Current Science* 80, 524–535.
- Gupta, P.K., Balyan, H.S., Edwards, K.J., Isaac, P., Korzun, V., Roder, M., Gautier, M.-F., Joudrier, P., Schlatter, A.R., Dubcovsky, J., De la Pena, R.C., Khairallah, M., Penner, G., Hayden, M.J., Sharp, P., Keller, B., Wang, R.C.C., Hardouin, J.P., Jack, P. and Leroy, P. (2002) Genetic mapping of 66 new microsatellite (SSR) loci in bread wheat. *Theoretical and Applied Genetics* 105, 413–422.
- Hackett, C.A. and Broadfoot, L.B. (2003) Effects of genotyping errors, missing values and segregation distortion in molecular marker data on the construction of linkage maps. *Heredity* 90, 33–38.
- Hackett, C.A. (2002) Statistical methods for QTL mapping in cereals. *Plant Molecular Biology* 48, 585–599.
- Haen, K.M., Lu, H.J., Friesen, T.L. and Faris, J.D. (2004) Genomic targeting and high-resolution mapping of the *Tsn1* gene in wheat. *Crop Science* 44, 951–962.
- Hearnden, P., Eckermann, P.J., McMichael, G.L., Hayden, M.J., Eglinton, J.K. and Chalmers, K.J. (2007) A genetic map of 1,000 SSR and DArT markers in a wide barley cross. *Theoretical and Applied Genetics* 115, 383–391.
- Hayden, M.J., Stephenson, P., Logojan, A.M., Khatkar, D., Rogers, C., Elsdon, J., Koebner, R.M.D., Snape, J.W. and Sharp, P.J. (2006) Development and genetic mapping of sequence-tagged microsatellites (STMs) in bread wheat (*Triticum aestivum* L.) *Theoretical and Applied Genetics* 113, 1271–1281.
- He, C.F., Sayed-Tabatabaei, B.E., Komatsuda, T. (2004) AFLP targeting of the 1-cM region conferring the *vrsl* gene for six-rowed spike in barley, *Hordeum vulgare* L. *Genome* 47, 1122–1129.
- Heun, M., Kennedy, A.E., Anderson, J.A., Lapitan, N.L.V., Sorrells, M.E. and Tanksley, S. D. (1991) Construction of a restriction fragment length polymorphism map for barley (*Hordeum vulgare*) *Genome* 34, 437–447.
- Hori, K., Kobayashi, T., Shimizu, A., Sato, K., Takeda, K. and Kawasaki, S. (2003) Efficient construction of high-density linkage map and its application to QTL analysis in barley. *Theoretical and Applied Genetics* 107, 806–813.
- Hori, K., Sato, K., Kobayashi, T. and Takeda, K. (2006) QTL analysis of fusarium head blight severity in recombinant inbred population derived from a cross between two-rowed barley varieties. *Breeding Science* 56, 25–30.

- Horsley, R.D., Schmierer, D., Maier, C., Kudrna, D., Urrea, C.A., Steffenson, B.J., Schwarz, P.B., Franckowiak, J.D., Green, M.J., Zhang, B. and Kleinhofs, A. (2006) Identification of QTLs associated with *Fusarium* head blight resistance in barley accession Clho 4196. *Crop Science* 46, 145–156.
- Hossain, K.G., Riera-Lizarazu, O., Kalavacharla, V., Vales, M.I., Maan, S.S. and Kianian, S. F. (2004) Radiation hybrid mapping of the species cytoplasm-specific (scs) gene in wheat. *Genetics* 168, 415–423.
- Huang, X., Zeller, F.J., Hsam, S.L.K., Wenzel, G. and Mohler, V. (2000) Chromosomal location of AFLP markers in common wheat utilizing nulli-tetrasomic stocks. *Genome* 43, 298–305.
- Islam, A.K.R.M., Shepherd, K.W. and Sparrow, D.H.B. (1981) Isolation and characterization of euplasmic wheat-barley chromosome addition lines. *Heredity* 46, 161–174.
- Jaccoud, D., Peng, K., Feinstein, D. and Kilian, A. (2001) Diversity Arrays: a solid state technology for sequence information independent genotyping. *Nucleic Acids Research* 29, e25.
- Jafary, H., Szabo, L.J. and Niks, R.E. (2006a) Innate nonhost immunity in barley to different heterologous rust fungi is controlled by sets of resistance genes with overlapping specificities. *Molecular Plant-Microbe Interactions* 19, 1270–1279.
- Jafary, H., Albertazi, G. and Niks, R.E. (2006b) Diversity of loci carrying genes for resistance of barley to heterologous rust species. In: Gullner, G. (ed.), *Proceedings of the non-specific and specific innate and acquired plant resistance symposium*, 31 August–3 September 2006, Budapest, Hungary. Plant Protection Institute of the Hungarian Academy of Sciences.
- Jansen, R.C. (1993) Interval mapping of multiple quantitative trait loci. *Genetics* 135, 205–211.
- Kalavacharia, V., Hossain, K., Gu, Y., Riera-Lizarazu, O., Vales, M.I., Bhamidimarri, S., Gonzales-Hernandez, J.L., Maan, S.S., Kianian, S. (2006) High resolution radiation hybrid map of wheat chromosome 1D. *Genetics* 106, 1089–1099.
- Karkousis, A., Barr, A.R., Kretschmer, J.M., Manning, S., Jefferies, S.P., Chalmers, K.J., Islam, A.K.M. and Langridge, P. (2003) Mapping and QTL analysis of the barley population Clipper \times Sahara. *Australian Journal of Agricultural Research* 54, 1137–1140.
- Kearsey, M.J. and Farquhar, A.G.L. (1998) QTL analysis in plants; where are we now? *Heredity* 80, 137–142.
- Kilian, A., Huttner, E., Wenzl, P., Jaccoud, D., Carling, J., Caig, V., Evers, M., Heller-Uszynska, K., Cayla, C., Patarapuwadol, S., Xia, L., Yang, S. and Thomson, B. (2005) The fast and the cheap: SNP and DArT-based whole genome profiling for crop improvement. In: Tuberosa, R., Phillips, R.L. and Gale, M. (eds.), *Proceedings of the International Congress "In the Wake of the Double Helix: From the Green Revolution to the Gene Revolution"*, 27–31 May 2003, Bologna, Italy. Avenue Media, pp. 443–461.
- Kleinhofs, A., Kilian, A., Saghai, M.M., Byashev, R.M., Hayes, P.M., Chen, F., Lapitan, N., Fenwick, A., Balkes, T.K., Kanazin, V., Ananiev, E., Dahleen, L., Kudrna, D., Bollinger, J., Knapp, S.J., Liu, B., Sorrels, M., Heun, M., Franckowiak, J.D., Hoffman, D., Skadsen, R. and Steffenson, B.J. (1993) A molecular isozyme and morphological map of barley (*Hordeum vulgare*) genome. *Theoretical and Applied Genetics* 86, 705–712.
- Korzun, V., Röder, M.S., Wendehake, K., Pasqualone, A., Lotti, C., Ganai, M.W. and Blanco, A. (1999) Integration of dinucleotide microsatellites from hexaploid bread wheat into a genetic linkage map of durum wheat. *Theoretical and Applied Genetics* 98, 1202–1207.
- Korzun, V., Malyshev, S., Voylovkov, A.V. and Borner, A. (2001) A genetic map of rye (*Secale cereale* L.) combining RFLP, isozyme, protein, microsatellite and gene loci. *Theoretical and Applied Genetics* 102, 709–717.

- Kumar, L.S. (1999) DNA markers in plant improvement: An overview. *Biotechnology Advances* 17, 143–182.
- Künzel, G., Korzun, L. and Meister, A. (2000) Cytologically integrated physical restriction fragment length polymorphism maps for the barley genome based on translocation break-points. *Genetics* 154, 397–412.
- Lande, R. and Thompson, R. (1990) Efficiency of marker-assisted selection in the improvement of quantitative traits. *Genetics* 124, 743–756.
- Lander, E.S. and Botstein, D. (1989) Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* 121, 185–199.
- Lander, E.S., Green, P., Abrahamson, J., Barlow, A., Daly, M.J., Lincoln, S.E. and Newburg, L. (1987) MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics* 1, 174–181.
- Langridge, P., Lagudah, E.S., Holton, T.A., Appels, R., Sharp, P.J. and Chalmers, K.J. (2001) Trends in genetic and genome analyses in wheat: a review. *Australian Journal of Agricultural Research* 52, 1043–1077.
- La Rota, M. and Sorrells, M.E. (2004) Comparative DNA sequence analysis of mapped wheat ESTs reveals the complexity of genome relationships between rice and wheat. *Functional and Integrative Genomics* 4, 34–46.
- Lehmensiek, A., Eckermann, P.J., Verbyla, A.P., Appels, R., Sutherland, M.W. and Daggard, G. (2005) Curation of wheat maps to improve map accuracy and QTL detection. *Australian Journal of Agricultural Research* 56, 1347–1354.
- Lehmensiek, A., Sutherland, M.W. and McNamara, R.B. (2008) The use of high resolution melting (HRM) to map single nucleotide polymorphism markers linked to a covered smut resistance gene in barley. *Theoretical and Applied Genetics* 117, 721–728.
- Liao, C.Y., Wu, P., Hu, B. and Yi, K.K. (2001) Effects of genetic background and environment of QTLs and epistasis for rice (*Oryza sativa* L.) panicle number. *Theoretical and Applied Genetics* 103, 104–111.
- Liew, M., Pryor, R., Palais, R., Meadows, C., Erali, M., Lyon, E. and Wittwer, C. (2004) Genotyping of single-nucleotide polymorphisms by high-resolution melting of small amplicons. *Clinical Chemistry* 50, 1156–1164.
- Li, D., Barclay, I., Jose, K., Stefanova, K. and Appels, R. (2008) A mutation at the Ala122 position of acetohydroxyacid synthase (AHAS) located on chromosome 6D of wheat: improved resistance to imidazolinone and a faster assay for marker assisted selection. *Molecular Breeding*, DOI 10.1007/s11032-008-9168-4.
- Li, H., Zhou, M., Mendham, N.J., Kilian, A., Wenzl, P., Huttner, E. and Vaillancourt, R.E. (2005) Mapping and QTL analysis of the barley population TX9428 × Franklin. In *Proceedings of the 12th Australian Barley Technical Symposium: 11–14 Sept Hobart, Australia*.
- Li, H., Vaillancourt, R.E., Zhou, M., Mendham, N.J., Wenzl, P., Huttner, E. and Kilian, A. (2006) Large-scale segregation distortion in barley revealed by a genetic map based on DArT, AFLP and SSR markers. In *Proceeding of the 13th Australasian Plant Breeding Conference: 18–21 April Christchurch, New Zealand*.
- Ling, H.Q., Zhu, Y. and Keller, B. (2003) High-resolution mapping of the leaf rust disease resistance gene *Lr1* in wheat and characterization of BAC clones from the *Lr1* locus. *Theoretical and Applied Genetics* 106, 875–882.
- Liu, Z.H., Anderson, J.A., Hu, J., Friesen, T.L., Rasmussen, J.B. and Faris, J.D. (2005) A wheat intervarietal genetic linkage map based on microsatellite and target region amplified polymorphism markers and its utility for detecting quantitative trait loci. *Theoretical and Applied Genetics* 111, 782–794.
- Lu, H.J., Fellers, J.P., Friesen, T.L., Meinhardt, S.W. and Faris, J.D. (2006) Genomic analysis and marker development for the *Tsn1* locus in wheat using bin-mapped ESTs and flanking BAC contigs. *Theoretical and Applied Genetics* 112, 1132–1142.

- Lynch, M. and Walsh, B. (1998) *Genetics and Analysis of Quantitative Traits*. Sinauer Associates, Sunderland.
- Ma, W., Appels, R., Bekes, F., Larroque, O.R., Morell, M.K. and Gale, K.R. (2005) Genetic characterisation of dough rheological properties in a wheat doubled haploid population: additive genetic effects and epistatic interactions. *Theoretical and Applied Genetics* 111, 410–422.
- Mago, R., Miah, H., Lawrence, G.J., Wellings, C.R., Spielmeier, W., Bariana, H.S., McIntosh, R.A., Pryor, A.J. and Ellis, J.G. (2005) High-resolution mapping and mutation analysis separate the rust resistance genes Sr31, Lr26 and Yr9 on the short arm of rye chromosome 1. *Theoretical and Applied Genetics* 112, 41–50.
- Manly, K.F., Cudmore, R.H.J. and Meer, J.M. (2001) Map Manager QTX, cross-platform software for genetic mapping. *Mammalian Genome* 12, 930–932.
- Mano, Y., Kawasaki, S., Takaiwa, F. and Komatsuda, T. (2001) Construction of a genetic map of barley (*Hordeum vulgare* L.) cross 'Azumamugi' × 'Kanto Nakate Gold' using a simple and efficient amplified fragment-length polymorphism system. *Genome* 44, 284–292.
- Marcel, T.C., Varshney, R.K., Barbieri, M., Jafary, H., de Kock, M.J.D., Graner, A. and Niks, R.E. (2007) A high-density consensus map of barley to compare the distribution of QTLs for partial resistance to *Puccinia hordei* and of defence gene homologues. *Theoretical and Applied Genetics* 114, 487–500.
- Martin, A., Rubiales, D. and Cabrera, A. (1999) A Fertile Amphiploid between a Wild Barley (*Hordeum chilense*) and Crested Wheatgrass (*Agropyron cristatum*). *International Journal of Plant Sciences* 160, 783–786.
- Masojć, P., Myśków, B. and Milczarski, P. (2001) Extending a RFLP-based genetic map of rye using random amplified polymorphic DNA (RAPD) and isozyme markers. *Theoretical and Applied Genetics* 102, 1273–1279.
- McCartney, C.A., Somers, D.J., Humphreys, D.G., Lukow, O., Ames, N., Noll, J., Cloutier, S. and McCallum, B.D. (2005) Mapping quantitative trait loci controlling agronomic traits in the spring wheat cross RL4452 × 'AC Domain'. *Genome* 48, 870–883.
- McFadden, H.G., Lehmensiek, A. and Lagudah, E.S. (2006) Resistance gene analogues of wheat: molecular genetic analysis of ESTs. *Theoretical and Applied Genetics* 113, 987–1002.
- Melchinger, A.E., Utz, H.F. and Schon, C.C. (1998) Quantitative trait locus mapping using different testers and independent population samples in maize reveals low power of QTL detection and large bias of QTL effects. *Genetics* 149, 383–403.
- Mesfin, A., Smith, K.P., Dill-Macky, R., Evans, C.K., Waugh, R., Gustus, C.D., Muehlbauer, G.J. (2003) Quantitative trait loci for Fusarium head blight resistance in barley detected in a two-rowed by six-rowed population. *Crop Science* 43, 307–318.
- Mester, D., Ronin, Y., Minkov, D., Nevo, E. and Korol, A. (2003a) Constructing large scale genetic maps using evolutionary strategy algorithm. *Genetics* 165, 2269–2282.
- Mester, D.I., Ronin, Y.I., Hu, Y., Nevo, E. and Korol, A. (2003b) Efficient multipoint mapping: Making use of dominant repulsion-phase markers. *Theoretical and Applied Genetics* 107, 1102–1112.
- Michelmore, R.W., Paran, I. and Kesseli, R. (1991) Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. *Proceedings of the National Academy of Sciences of the United States of America* 88, 9829–9832.
- Mingeot, D. and Jacquemin, J.M. (1999) Mapping of RFLP probes characterized for their polymorphism on wheat. *Theoretical and Applied Genetics* 98, 1132–1137.
- Montgomery, J., Wittwer, C.T., Palais, R. and Luming, Z. (2007) Simultaneous mutation scanning and genotyping by high-resolution DNA melting analysis. *Nature* 2, 59–66.

- Miftahudin, T., Chikmawati, T., Ross, K., Scoles, G.J. and Gustafson, J.P. (2005) Targeting the aluminum tolerance gene *Alt3* region in rye, using rice/rye micro-colinearity. *Theoretical and Applied Genetics* 110, 906–913.
- Nachit, M.M., Elouafi, I., Pagnotta, M.A., El Saleh, A., Iacono, E., Labhilili, M., Asbati, A., Azrak, M., Hazzam, H., Benscher, D., Khairallah, M., Ribaut, J.M., Tanzarella, O.A., Porceddu, E. and Sorrells, M.E. (2001) Molecular linkage map for an intraspecific recombinant inbred population of durum wheat (*Triticum turgidum* L. var. *durum*). *Theoretical and Applied Genetics* 102, 177–186.
- Nelson, J.C., Sorrells, M.E., Van Deynze, A.E., Hai Lu, Y., Atkinson, M., Bernard, M., Leroy, P., Faris, J.D. and Anderson, J.A. (1995) Molecular mapping of wheat: Major genes and rearrangements in homoeologous groups 4, 5, and 7. *Genetics* 141, 721–731.
- Paillard, S., Schnurbusch, T., Winzeler, M., Messmer, M., Sourdille, P., Abderhalden, O., Keller, B. and Schechermayr, G. (2003) An integrative genetic linkage map of winter wheat (*Triticum aestivum* L.). *Theoretical and Applied Genetics* 107, 1235–1242.
- Parker, G.D., Chalmers, K.J., Rathjen, A.J. and Langridge, P. (1999) Mapping loci associated with milling yield in wheat (*Triticum aestivum* L.). *Molecular Breeding* 5, 561–568.
- Peleman, J.D., Wye, C., Zethof, J., Sorensen, A.P., Verbakel, H., van Oeveren, J., Gerats, T. and van der Voort, J.R. (2005) Quantitative trait locus (QTL) isogenic recombinant analysis: A method for high-resolution mapping of QTL within a single population. *Genetics* 171, 1341–1352.
- Pellio, B., Streng, S., Bauer, E., Stein, N., Perovic, D., Schiemann, A., Friedt, W., Ordon, F. and Graner, A. (2005) High-resolution mapping of the Rym4/Rym5 locus conferring resistance to the barley yellow mosaic virus complex (BaMMV, BaYMV, BaYMV-2) in barley (*Hordeum vulgare* ssp. *vulgare* L.). *Theoretical and Applied Genetics* 110, 283–293.
- Pena, R.J. (2002) Wheat for bread and other foods. In: Curtis, B.C., Rajaram, S. and Gomez Macpherson, H. (eds.), *Bread Wheat: Improvement and Production*. Food and Agriculture Organization of the United Nations, Rome.
- Peng, J., Richards, D.E., Hartley, N.M., Murphy, G.P., Devos, K.M., Flinthan, J.E., Beales, J., Fish, L.J., Worland, A.J., Pelica, F., Sudhakar, D., Christou, P., Snape, J.W., Gale, M.D. and Harberd, N.P. (1999) ‘Green revolution’ genes encode mutant gibberellin responses modulators. *Nature* 400, 256–261.
- Prasad, M., Kumar, N., Kulwal, P.L., Roder, M.S., Balyan, H.S., Dhaliwal, H.S. and Gupta, P.K. (2003) QTL analysis for grain protein content using SSR markers and validation studies using NILs in bread wheat. *Theoretical and Applied Genetics* 106, 659–667.
- Qi, L., Echalié, B., Friebe, B. and Gill, B.S. (2003) Molecular characterization of a set of wheat deletion stocks for use in chromosome bin mapping of ESTs. *Functional and Integrative Genomics* 3, 39–55.
- Qi, X., Stam, P. and Lindhout, P. (1998) Use of locus-specific AFLP markers to construct a high density-molecular map in barley. *Theoretical and Applied Genetics* 96, 376–384.
- Quarrie, S.A., Steed, A., Calestani, C., Semikhodskii, A., Lebreton, C., Chinoy, C., Steele, N., Pljevljakusić, D., Waterman, E., Weyen, J., Schondelmaier, J., Habash, D.Z., Farmer, P., Saker, L., Clarkson, D.T., Abugalieva, A., Yessimbekova, M., Turuspekov, Y., Abugalieva, S., Tuberosa, R., Sanguineti, M.C., Hollington, P.A., Aragués, R., Royo, A. and Dodig, D. (2005) A high-density genetic map of hexaploid wheat (*Triticum aestivum* L.) from the cross Chinese Spring × SQ1 and its use to compare QTLs for grain yield across a range of environments. *Theoretical and Applied Genetics* 110, 865–880.
- Raman, H., Platz, G.J., Chalmers, K.J., Raman, R., Read, B.J., Barr, A.R. and Moody, D.B. (2003) Mapping of genomic regions associated with net form of net blotch resistance in barley. *Australian Journal of Agricultural Research* 54, 1359–1367.
- Ramsay, L., Macaulay, M., degli Ivanisovich, S., MacLean, K., Cardle, L., Fuller, J., Edwards, K.J., Turesson, S., Morgante, M., Massari, A., Maestri, E., Marmiroli, N.,

- Sjakste, T., Ganal, M., Powell, W. and Waugh, R. (2000) A simple sequence repeat-based linkage map of barley. *Genetics* 156, 1997–2005.
- Röder, M.S., Korzun, V., Wendehake, J., Plaschke, J., Tixier, M.H., Leroy, P. and Ganal, M. W. (1998) A microsatellite map of wheat. *Genetics* 149, 2007–2023.
- Rostoks, N., Mudie, S., Cardle, L., Russell, J., Ramsay, L., Booth, A., Svensson, J.T., Wanamaker, S.I., Walia, H., Rodriguez, E.M., Hedley, P.E., Liu, H., Morris, J., Close, T.J., Marshall, D.F. and Waugh, R. (2005a) Genome-wide SNP discovery and linkage analysis in barley based on genes responsive to abiotic stress. *Molecular and General Genomics* 274, 515–527.
- Rostoks, N., Borevitz, J.O., Hedley, P.E., Russell, J., Mudie, S., Morris, J., Cardle, L., Marshall, D.F. and Waugh, R. (2005b) Single-feature polymorphism discovery in the barley transcriptome. *Genome Biology* 6.
- Sears, E.R. (1954) The aneuploids of common wheat. *Missouri Agricultural Experiment Station Research Bulletin*, pp. 1–59.
- Sears, E.R. and Sears, L.M.S. (1978) The telocentric chromosomes of common wheat. Paper presented at the Proceedings of the 5th International Wheat Genetic Symposium, New Dehli.
- Shang, H.-Y., Wei, Y.-M., Wang, X.-R. and Zheng, Y.-L. (2006) Genetic diversity and phylogenetic relationships in the rye genus *Secale* L. (rye) based on *Secale cereale* microsatellite markers. *Genetics and Molecular Biology* 29, 685–691.
- Schnurbusch, T., Bossolini, E., Messmer, M. and Keller, B. (2004) Tagging and validation of a major quantitative trait locus for leaf rust resistance and leaf tip necrosis in winter wheat cultivar Forno. *Phytopathology* 94, 1036–1041.
- Somers, D.J., Fedak, G. and Savard, M. (2003) Molecular mapping of novel genes controlling Fusarium head blight resistance and deoxynivalenol accumulation in spring wheat. *Genome* 46, 555–564.
- Somers, D.J., Isaac, P. and Edwards, K. (2004) A high-density microsatellite consensus map for bread wheat (*Triticum aestivum* L.). *Theoretical and Applied Genetics* 109, 1105–1114.
- Song, Q.J., Shi, J.R., Singh, S., Fickus, E.W., Costa, J.M., Lewis, J., Gill, B.S., Ward, R. and Cregan, P.B. (2005) Development and mapping of microsatellite (SSR) markers in wheat. *Theoretical and Applied Genetics* 110, 550–560.
- Sourdille, P., Cadalen, T., Guyomarç'h, H., Snape, J.W., Perretant, M.R., Charmet, G., Boeuf, C., Bernard, S. and Bernard, M. (2003) An update of the Courtot-Chinese Spring intervarietal molecular marker linkage map for the QTL detection of agronomic traits in wheat. *Theoretical and Applied Genetics* 106, 530–538.
- Spielmeyer, W., Singh, R.P., McFadden, H., Wellings, C.R., Huerta-Espino, J., Kong, X., Appels, R. and Lagudah, E.S. (2008) Fine scale genetic and physical mapping using deletion mutants of Lr34/Yr18: A disease resistance locus effective against multiple pathogens in wheat. *Theoretical and Applied Genetics*, doi: 10.1007/s00122-007-0684-0.
- Stam, P. (1993) Construction of integrated genetic linkage maps by means of a new computer package: JoinMap. *The Plant Journal* 3, 739–744.
- Stein, N., Feuillet, C., Wicker, T., Schlagenhauf, E. and Keller, B. (2000) Subgenome chromosome walking in wheat: A 450-kb physical contig in *Triticum monococcum* L. spans the Lr10 resistance locus in hexaploid wheat (*Triticum aestivum* L.). *Proceedings of the National Academy of Sciences of the United States of America* 97, 13436–13441.
- Stein, N., Prasad, M., Scholz, U., Thiel, T., Zhang, H., Wolf, M., Kota, R., Varshney, R., Perovic, D., Grosse, I. and Graner, A. (2007) A 1,000-loci transcript map of the barley genome: new anchoring points for integrative grass genomics. *Theoretical and Applied Genetics* 114, 823–839.

- Sutton, T., Baumann, U., Hayes, J., Collins, N.C., Shi, B.J., Schnurbusch, T., Hay, A., Mayo, G., Pallotta, M., Tester, M. and Langridge, P. (2007) Boron-toxicity tolerance in barley arising from efflux transporter amplification. *Science* 318, 1446–1449.
- Szakacs, E. and Molnar-Lang, M. (2007) Development and molecular cytogenetic identification of new winter wheat-winter barley disomic addition lines. *Genome* 50, 43–50.
- Thangavelu, M., James, A.B., Bankier, A., Bryan, G.J., Dear, P.H. and Waugh, R. (2003) HAPPY mapping in a plant genome: reconstruction and analysis of a high resolution physical map of a 1.9 Mbp region of *Arabidopsis thaliana* chromosome 4. *Plant Biotechnology Journal* 1, 23–31.
- Taketa, S., Awayama, T., Amano, S., Sakurai, Y. and Ichii, M. (2006) High-resolution mapping of the nud locus controlling the naked caryopsis in barley. *Plant Breeding* 125, 337–342.
- Tanksley, S.D. (1993) Mapping polygenes. *Annual Review of Genetics* 27, 205–233.
- Teuscher, F. and Broman, K.W. (2007) Haplotype probabilities for multiple strain recombination inbred lines. *Genetics* 175, 1267–1274.
- Torada, A., Koike, M., Mochida, K. and Ogihara, Y. (2006) SSR-based linkage map with new markers using an intraspecific population of common wheat. *Theoretical and Applied Genetics* 112, 1042–1051.
- United States Department of Agriculture (2006) Wheat: World Markets and Trade <<http://www.fas.usda.gov/grain/>>.
- Van Os, H., Andrzejewski, S., Bakker, E., Barrena, I., Bryan, G.J., Caromel, B., Ghareeb, B., Isidore, E., de Jong, W., van Koert, P., Lefebvre, V., Milbourne, D., Ritter, E., Rouppe van der Voort, J.N.A.M., Rousselle-Bourgeois, F., van Vliet, J., Waugh, R., Visser, R.G.F., Bakker, J. and Van Eck, H.J. (2006) Construction of a 10,000-marker ultradense genetic recombination map of potato: Providing a framework for accelerated gene isolation and a genome wide physical map. *Genetics* 173, 1075–1087.
- Van Os, H., Stam, P., Visser, R.G.F. and Van Eck, H.J. (2005a) RECORD: a novel method for ordering loci on a genetic linkage map. *Theoretical and Applied Genetics* 112, 30–40.
- Van Os, H., Stam, P., Visser, R.G.F. and Van Eck, H. (2005b) SMOOTH: a statistical method for successful removal of genotyping errors from high-density genetic linkage data. *Theoretical and Applied Genetics* 112, 187–194.
- Varshney, R.K., Grosse, I., Hähnel, U., Siefken, R., Prasad, M., Stein, N., Langridge, P., Altschmied, L. and Graner, A. (2006) Genetic mapping and BAC assignment of EST-derived SSR markers shows non-uniform distribution of genes in the barley genome. *Theoretical and Applied Genetics* 113, 239–250.
- Varshney, R.K., Marcel, T.C., Ramsay, L., Russell, J., Röder, M.S., Stein, N., Waugh, R., Langridge, P., Niks, R.E. and Graner, A. (2007) A high density barley micro-satellite consensus map with 775 SSR loci. *Theoretical and Applied Genetics* 114, 1091–1103.
- Verbyla, A.P., Cullis, B.R. and Thompson, R. (2007) The analysis of QTL by simultaneous use of the full linkage map. *Theoretical and Applied Genetics* 116, 95–111.
- Vos, P., Hogers, R., Beleecker, M., Reijans, M., Van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M. and Zabeau, M. (1995) AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research* 23, 4407–4414.
- Wang, J., Raman, H., Zhou, M., Ryan, P.R., Delhaize, E., Hebb, D.M., Coombes, N. and Mendham, N. (2007) High-resolution mapping of the Alp locus and identification of a candidate gene HvMATE controlling aluminium tolerance in barley (*Hordeum vulgare* L.). *Theoretical and Applied Genetics* 115, 265–276.
- Wardrop, J., Snape, J., Powell, W. and Machray, G.C. (2002) Constructing plant radiation hybrid panels. *The Plant Journal* 31, 223–228.

- Wardrop, J., Fuller, J., Powell, W. and Machray, G.C. (2004) Exploiting plant somatic radiation hybrids for physical mapping of expressed sequenced tags. *Theoretical and Applied Genetics* 108, 343–348.
- Wenzl, P., Carling, J., Kudrna, D., Jaccoud, D., Huttner, E., Kleinhofs, A. and Kilian, A. (2004) Diversity arrays technology (DArT) for whole-genome profiling of barley. *Proceedings of the National Academy of Sciences of the United States of America* 101, 9915–9920.
- Wenzl, P., Li, H., Carling, J., Zhou, M., Raman, H., Paul, E., Hearnden, P., Maier, C., Xia, L., Caig, V., Ovesna, J., Cakir, M., Poulsen, D., Wang, J., Raman, R., Smith, K.P., Meuhlbauer, G.J., Chalmers, K.J., Kleinhofs, A., Huttner, E. and Kilian, A. (2006) A high-density consensus map of barley linking DArT markers to SSR, RFLP and STS loci and agricultural traits. *BMC Genomics* 7, 206–227.
- Wenzl, P., Raman, H., Wang, J., Zhou, M., Huttner, E., Kilian, A. (2007a) A DArT platform for quantitative bulked segregant analysis. *BMC Genomics* 8, 196–206.
- Wenzl, P., Catizone, I., Thomson, B., Huttner, E., Mantovani, P., Maccaferri, M., Corneti, S., De Ambrogio, E., Sanguineti, M.C., Tuberosa, R. and Kilian, A. (2007b) A DArT platform for high-throughput profiling of durum wheat. *Plant and Animal Genomes XIV Conference*, 13–17 Jan, San Diego, CA, USA.
- Williams, J.G.K., Kubelik, A.R., Livak, K.J., Rafalski, J.A. and Tingey, S.V. (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research* 18, 6531–6535.
- Willmore, K.L., Eckermann, P., Varshney, R.K., Graner, A., Langridge, P., Pallotta, M., Cheong, J. and Williams, K.J. (2006) New eSSR and gSSR markers added to Australian barley maps. *Australian Journal of Agricultural Research* 57, 953–959.
- Wu, J., Jenkins, J., Zhu, J., McCarty, J. and Watson, C. (2003) Monte Carlo simulations on marker grouping and ordering. *Theoretical and Applied Genetics* 107, 568–573.
- Xing, Y.Z., Tan, Y.F., Hua, J.P., Sun, X.L. and Xu, C.G. (2002) Characterization of the main effects, epistatic effects and their environmental interactions of QTLs on the genetic basis of yield traits in rice. *Theoretical and Applied Genetics* 105, 248–257.
- Yamamoto, M. and Mukai, Y. (2005) High-resolution physical mapping of the secalin-1 locus of rye on extended DNA fibers. *Cytogenetic and Genome Research* 109, 79–82.
- Yandell, B.S., Mehta, T., Banerjee, S., Shriner, D., Venkataraman, R., Moon, J.Y., Neely, W., Wu, H. von Smith, R. and Yi, N. (2007) R/qtlbim: QTL with Bayesian Interval Mapping in experimental crosses. *Bioinformatics* 23, 641–643.
- Yang, H.Y., You, A.Q., Yang, Z.F., Zhang, F., He, R.F., Zhu, L.L. and He, G. (2004) High-resolution genetic mapping at the Bph15 locus for brown planthopper resistance in rice (*Oryza sativa* L.). *Theoretical and Applied Genetics* 110, 182–191.
- Yap, I., Schneider, D., Kleinberg, J., Matthews, D., Cartinhour, S. and McCouch, S. (2003) A graph-theoretic approach to comparing and integrating genetic, physical and sequence-based maps. *Genetics* 165, 2235–2247.
- Yokota, H., Johnson, F., Lu, H., Robinson, M., Belu, A.M., Garrison, M.D., Ratner, B.D., Trask, B.J., Miller, D.L. (1997) A new method for straightening DNA molecules for optical retraction mapping. *Nucleic Acids Research* 25, 1064–1070.
- Young, N.D. (1996) QTL mapping and quantitative disease resistance in plants. *Annual Review of Phytopathology* 34, 479–501.
- Yu, S.B., Li, J.X., Xu, C.G., Tan, Y.F., Gao, Y.J., Li, X.H., Zhang, Q. and Saghai Maroof, M.A. (1997) Importance of epistasis as the genetic basis of heterosis in an elite rice hybrid. *Proceedings of the National Academy of Sciences of the United States of America* 94, 9226–9231.
- Zeng, Z.-B. (1994) Precision mapping of quantitative trait loci. *Genetics* 136, 1457–1468.

- Zhou, S., Bechner, M.C., Place, M., Churas, C.P., Pape, L., Leong, S.A., Runnheim, R., Forrest, D.K., Goldstein, S., Livny, M. and Schwartz, D.C. (2007) Validation of rice genome sequence by optical mapping. *BMC Genomics* 8, 278–296.
- Zhuang, J.-Y., Fan, Y.-Y., Rao, Z.-M., Wu, J.-L., Xia, Y.-W. and Zheng, K.-L. (2002) Analysis of additive effects and additive-by-additive epistatic effects of QTLs for yield traits in a recombinant inbred line population of rice. *Theoretical and Applied Genetics* 105, 1137–1145.

Chapter 8

Early Stages of Meiosis in Wheat- and the Role of *Ph1*

Graham Moore

Abstract Studies have revealed a number of loci which control chromosome pairing and recombination in wheat. Exploitation of such loci could have a major impact on breeding strategies. The review summarises our current knowledge of this process, with particular emphasis on the most extensively studied chromosome pairing locus to date, *Ph1* (*Pairing homoeologous 1*). The *Ph1* story to date has revealed that chromatin remodeling at the start of meiosis and the control of early stages of meiotic prophase I are both important.

8.1 The Introduction

Meiosis is a cell division process that ensures gametes carry the correct number of chromosomes, without a doubling of chromosome number (Fig. 8.1). It is therefore central to the life cycle of all sexual eukaryotic organisms. Such species carry two copies (homologues) of each chromosome that are separated during meiosis into different gametes. Each homologue is derived from one of its parents. The original chromosome number is only restored after fertilisation. Before meiosis, each homologue is replicated, forming two sister chromatids that remain linked together. At the start of meiosis, each chromosome (composed of two sister chromatids) must recognise its homologue from amongst all the chromosomes present in the nucleus. Having recognised its correct partner, the two homologues must then intimately align along their entire lengths. As part of this alignment, a proteinaceous structure known as the synaptonemal complex (SC) is assembled between the homologues, in a process called synapsis, reviewed by Zickler and Kleckner (1999). Within this structure, meiotic recombination (the exchange of DNA strands) is completed, resulting in the crossover formation between the DNA strands of the homologues or parental chromosomes, thereby reshuffling genetic information. These crossovers, together with sister chromatid cohesion, provide

G. Moore (✉)

John Innes Centre, Colney, Norwich. NR4 7UH, UK

e-mail: graham.moore@bbsrc.ac.uk

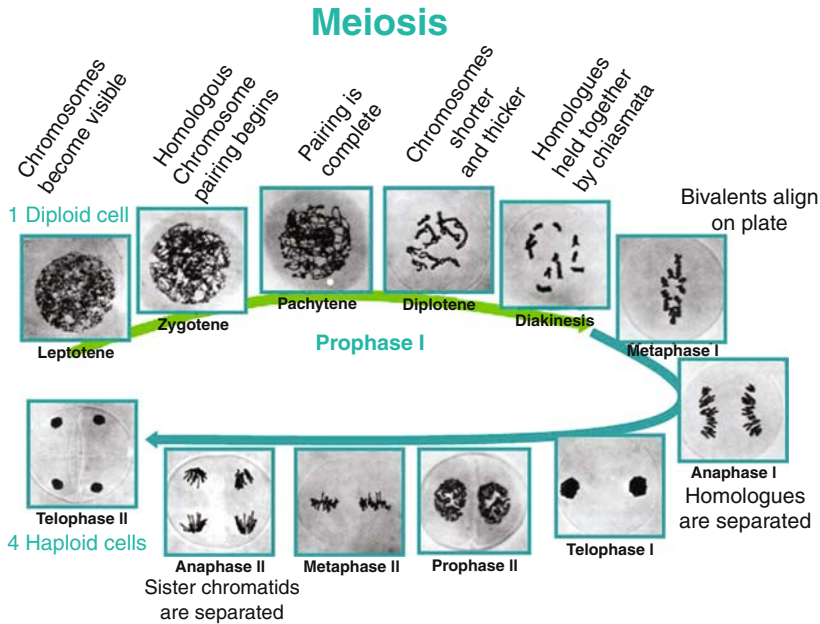


Fig. 8.1 Meiosis. Cytological appearance of chromosomes at the various defined stages of meiosis (See Color Insert)

physical links (visualised as chiasma) that hold the homologues together after the SC is disassembled and allow the correct orientation of homologues on the first meiotic spindle. The homologues are then correctly segregated during meiosis I. A second division occurs in which the sister chromatids are segregated so that each gamete carries only a single copy of each chromosome (Page and Hawley 2003).

While over the past two decades comprehensive research investigating meiosis has been undertaken in model organisms such as yeast and *Arabidopsis*, it has not been until recently that studies have been completed in more complex genomes like hexaploid (bread) wheat. Identifying the corresponding components of key recombination and SC wheat proteins previously characterised in yeast and other models has begun to unfold with several key candidates having been reported (Sutton et al. 2003; Crismani et al. 2006; Boden et al. 2007; Lloyd et al. 2007). The approach identifies key candidates which can be subject to modification and their effects on enhancement of increased, or altered distribution of recombination then assessed. Of course, the ultimate goal of such studies is the enhancement of breeding.

8.2 Chromosome Sorting for Meiosis

How are chromosomes sorted into partners and how are chromosomes recognised during the early stages of meiosis? Studies of yeast (*Schizosaccharomyces pombe* and *Saccharomyces cerevisiae*), mammals, rye, maize and wheat show

that telomeres of chromosomes aggregate on the nuclear envelope forming a telomere cluster or bouquet during early meiotic prophase I (Tresses-Sticken et al. 1999; Chikashige et al. 1997; Niwa et al. 2000; Mikhailova et al. 2001; Martinez-Perez et al. 1999; Bass et al. 1997). The proposal is that this structure facilitates in some way the sorting of homologous chromosomes into pairs. Synaptonemal complex formation initiates near the telomeres. Deletion of the telomere region of one of the homologues is generally sufficient to reduce or eliminate subsequent pairing between these chromosomes at metaphase, which may reflect a failure in initiation of synapsis (Curtis et al. 1991; Lukaszewski 1997). Thus, homologous chromosomes are tethered at a region close to their telomeres during early prophase I and then intimately aligned along their length as prophase I pairing proceeds. Synapsis of the chromosomes also is initiated from the telomere regions (Zickler and Kleckner 1999).

8.3 Recombination- Factors Affecting its Distribution

Recombination has a major influence on breeding. There are at least three constraints that affect crossover distribution. Each of the synapsed pairs of chromosomes must possess at least one crossover, this is known as the obligatory crossover. Crossovers are not independent of each other as the presence of a crossover inhibits the formation of a new crossover nearby on the chromosome and this inhibition decreases with distance. This process is known as crossover interference (Jones 1984). Crossovers are not uniformly distributed along chromosomes, some regions, known as hotspots, display much higher frequencies of crossing over than others. This last point is very important for breeding programs since it implies that the distribution of crossovers with respect to the location of genes will affect the efficiency of the breeding process. Many species with large genomes (for example wheat, barley, *Lolium*, *Festuca* and rye) exhibit an uneven distribution of recombination along their chromosomes with the proximal chromosome regions being highly suppressed in recombinational activity (Akhunov et al. 2003). These species such as wheat, barley and rye with large chromosomes possess extensive numbers of similar repetitive sequences throughout the genome. In fact most of the intervening DNA between genes in the large genomes is composed of these repetitive elements or retroelements. It could be thought that the presence of highly similar repetitive sequences might compromise the ability of homologous chromosomes to recognise each other in early meiotic prophase I. However comparative sequencing of cereals and grasses has revealed that although the genes are largely conserved in equivalent regions in related species, the repetitive elements are not conserved. Sequence comparisons of the repetitive content of wheat chromosomes with their diploid progenitors also revealed lack of conservation of repeat sequences in the intervening regions suggesting a rapid turnover (Paux et al. 2006; Scherrer et al. 2005). In the most extreme case of

maize, the intervening regions between two varieties can be markedly different (Fu and Dooner 2002). Thus recombination between the chromosomes of related species or in the case of maize, between the chromosomes of two varieties is going to be confined largely to the homologous regions or the genic regions. Hence the distribution of genes along a chromosome with respect to the repetitive elements will influence the distribution of recombination as they are the key points of high homology between chromosomes. The uneven distribution of recombination along the chromosomes of wheat, barley and rye may imply an uneven distribution of genes. For many of these species, the genome sequence is not yet available so it is unclear whether the regions lacking recombinational activity also lack genes. However early comparative mapping work revealed that whole rice chromosomes (particularly linkage groups 8, 10, 11 and 12) were equivalent to the proximal regions of chromosomes in wheat, rye, barley and *Lolium* (Moore 1995; King et al. 2007). These rice chromosomes contain thousands of genes and therefore by implication, so must the equivalent regions in wheat. Studies exploiting barley deletions also indicated that rice linkage groups could be allocated to cytologically defined regions (Korzun and Kunzel 1996; Kunzel et al. 2000). These studies suggested the presence of significant numbers of genes in these regions lacking recombinational activity. Recent studies exploiting comparative mapping in *Lolium/Festuca*, wheat and rice have provided further evidence to support the suggestion that there are a significant percentage of genes in regions of poor recombination (King et al. 2007). Thus the genes in these regions of the wheat genome are likely to remain as a tight linkage group during breeding. An important goal of meiotic studies in wheat is therefore to alter the distribution of recombination.

8.4 Polyploids

Polyploids possess two or more sets of related chromosomes as a result of either the doubling of chromosomes following sexual hybridisation within the same species (autopolyploid), or between closely related species containing related but not completely homologous (homoeologous-) genomes (allopolyploid). More than 70% of flowering plants are thought to be polyploidy (Materson 2003; Bowers et al. 2003). Moreover genome analysis reveals that those species which were originally thought of as diploid (including rice and *Arabidopsis*), contain whole genome duplications. The hexaploid wheat genome with its diploid set of 42 chromosomes is composed of three ancestral genomes, A, B and D. The 42 chromosomes can be divided into seven groups of six chromosomes (two chromosomes from each of the ancestral genomes). Others such as maize are cryptic allotetraploids; that is following polyploidisation, genome rearrangements have partially concealed their polyploid origin. Therefore in allopolyploids, homologues are distinguished from related chromosomes (homoeologues) with similar gene content and order but are divergent in

repetitive DNA content. Hexaploid and tetraploid wheat both behave as diploids at meiosis, with regular pairing at metaphase I. Tetraploid wheat possesses just the A and B homoeologous genomes, and lacks the D genome. In hexaploid wheat chromosome 1A has a similar gene order to 1B and 1D, but must pair at meiosis with its homologue 1A. The overall efficiency and accuracy of the mechanisms used to achieve this diploidisation has a profound influence on the fertility of wheat, and this is clearly of major importance for the success of a crop, such as wheat in which seed is the economic product. This diploidisation mechanism also has an effect on the ability to exploit wild species in breeding programmes, of which many carry interesting characteristics that would be useful to exploit. The sexual hybridisation between wheat and a wild species generally produces an interspecific hybrid containing a haploid set of polyploid and wild-relative chromosomes. The exploitation of alien variation during breeding of wheat benefited from an understanding of the genome relationships of wheat and its wild diploid and polyploid relatives. As a result, the genomic relationships within the *Triticeae* have been extensively studied. By 1952, the work of Kihara and colleagues had established many of the genomic relationships of the diploid and polyploid *Triticum* and *Aegilops* species on the basis of chromosome pairing in hybrids. These studies have continued and have accurately classified several translocations. Chromosome pairing was assessed by squashing meiocytes at metaphase I and classifying the specific chromosome configurations at this stage. The association of chromosomes implied structural relationship between chromosomes (including genes). However -in many cases, there is low level of pairing and recombination between crop and wild relative chromosomes even though molecular studies have subsequently suggested that the chromosomes are structurally similar. The factors which affect the ability to pair and recombine these chromosomes are often those involved in controlling chromosome pairing between the related chromosomes within a polyploid such as wheat itself.

8.5 Chromosome Pairing Loci

The diploid chromosome pairing behaviour which restricts pairing to homologues rather than homoeologues in wheat is under genetic control. A single locus, *Ph1*, (*Pairing homoeologous 1*) on the long arm of chromosome 5B has a major controlling effect (Riley and Chapman 1958). However, other loci that affect pairing in wheat have also been identified including *Ph2* on 3DS as well as loci on chromosomes 5AL, 5DL, 5AS, 3AL, 3BL and 3DL (Feldman 1993). Even so, these other loci do not compensate for the absence of the *Ph1* locus. Compared to the characterisation of other pairing loci, *Ph1* has been the most extensively studied to date. Therefore understanding the cell biological and molecular basis of other loci such as *Ph2* is not so advanced. What is known is that *Ph2* appears to affect chromosome synapsis (Martinez et al. 2001), and

based on a comparative study between rice and wheat, a number of candidate genes have been linked to this region which may have a role in this process (Dong et al. 2002; Sutton et al. 2003).

In the presence of the *Ph1* locus (that is, two copies of *Ph1*), chromosome pairing in tetraploid and hexaploid wheat during meiosis is largely restricted to homologous chromosomes, but in deletion mutants lacking *Ph1*, increasing aberrant pairing can be observed at metaphase I with each generation of the mutant (Sears 1977; Roberts et al. 1999) (Fig. 8.2). With each generation, wheat lacking *Ph1* accumulates extensive rearrangements such that it eventually becomes sterile (Sanchez-Moran et al. 2001). Therefore the *Ph1* deletion stocks are maintained in a heterozygous condition in order to avoid homoeologous pairing between wheat chromosomes.

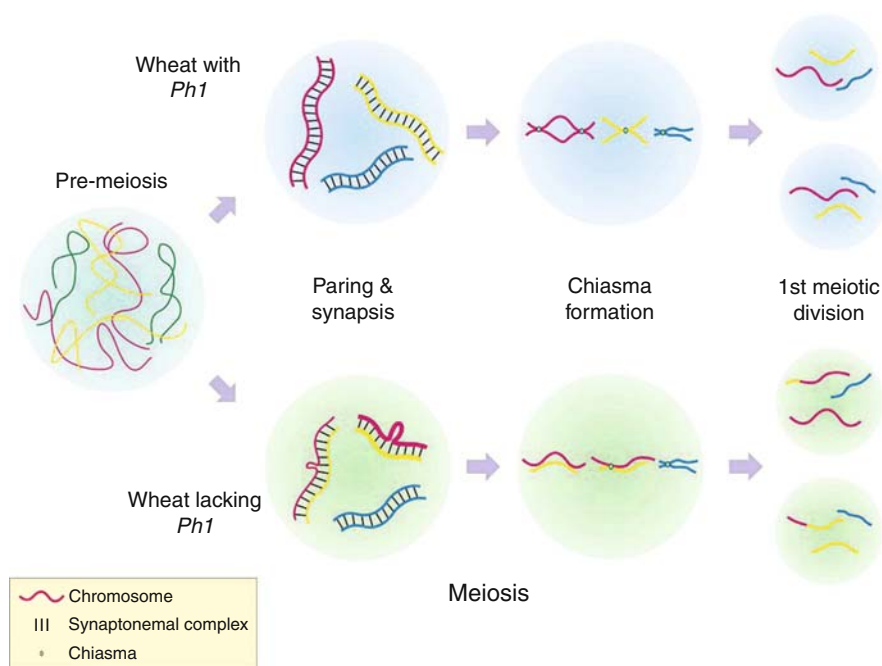


Fig. 8.2 A schematic representation of chromosome pairing in the presence and absence of the *Ph1* locus (See Color Insert)

8.6 The *Ph1* Locus

So how does *Ph1* work? Early cytogenetic studies revealed that in hybrids between wheat and related species, the addition of B chromosomes compensated for the absence of the *Ph1* locus (Dover and Riley 1972). Their presence

reduced the level of homoeologous pairing at metaphase I as observed in squashed meiocyte preparations. Later studies revealed that B chromosomes reduced the synchronized replication of related heterochromatin in the (maize) genome (Pryor et al. 1980). The implication of these early observations was that the *Ph1* locus affected replication and heterochromatin behaviour in some way. With the advent of the ability to visualise chromosome synapsis in hexaploid wheat, Holm showed that in the presence of *Ph1*, only five out of 42 chromosomes engage in multiple associations at the start of synapsis (Holm 1986, 1988; Holm and Wang 1988; Martinez et al. 2001). These multiple associations are resolved in later prophase I. However in the absence of *Ph1*, some 19 chromosomes out of 42 engage in multiple associations at the start of pairing and many of these associations remain unresolved even in late prophase I. With the advent of Restriction Fragment Length Polymorphisms (RFLPs) and the construction of genetic maps of hexaploid wheat, Dvorak and colleagues found that in the absence of *Ph1*, recombination occurred between a pair of wheat chromosomes composed of combinations of homoeologous and homologous segments, but in the presence of *Ph1*, recombination was restricted to homologous segments (Dubcovsky et al. 1995; Luo et al. 1996). Thus the *Ph1* locus affected the stringency of the recombinational process during meiosis as revealed by Dvorak's data exploiting genetic mapping and synapsis as revealed by Holm's data showing the failure of multivalents to resolve at late prophase I. However the effect of B chromosomes on wheat meiosis implied that *Ph1* also had an effect on heterochromatin and replication as well.

As described above, squashed preparations of meiocytes are routinely used to score chromosome pairing at metaphase I in wheat where they provide an accurate score of the level of pairing, as the pairing at this stage is unaffected by the procedures used to generate the preparations. However such preparations have been also used to study early stages in meiosis and even in premeiosis. The problem with using such "squashed" preparations to study these stages is that the process of squashing disrupts the chromosome structures and initial interactions between chromosomes. Moreover it is difficult to distinguish cell types such meiocytes, tapetal and cell wall cells after squashing particularly in premeiotic stages. Thus the exploitation of such preparations has often led to contradictory observations with one study initially reporting altered chromosome structure after squashing only for a later study to indicate that there is no altered structure (Mikhailova et al. 1998; Maestra et al. 2002; Kopecky et al. 2007). Put simply- if a block of stone is dropped on your head and you are flattened, your head will pair with your feet when of course it should not normally do this! The advent of cell biological approaches in the last 10 years solved these problems and has enabled the effects of *Ph1* on heterochromatin to be studied in more detail. Cell biology studies exploiting *in situ* hybridisation, anther sectioning and confocal microscopy enables the behaviour of chromosomes to be analysed in 3D in all the cells in an anther. Such 3-D projections can then be rotated so that the process can be visualized more easily from any angle. Examples of whole sections can be seen for each stage described below on the

following websites (<http://www.jic.ac.uk/staff/graham-moore/> and then go to meiosis gallery or <http://www.jic.ac.uk/staff/peter-shaw/meiosisGallery.html>). Cell biological studies showed that wheat and rice chromosomes can also associate in pairs prior to meiosis during floral development by their centromeres (Martinez-Perez et al. 2000; Martinez-Perez et al. 2001, 2003; Prieto et al. 2004b). *In situ* hybridisation demonstrated that in hexaploid wheat prior to meiosis its 42 centromeres are visualised as 21 sites, thus implying pairing of the centromeres. In tetraploid wheat prior to meiosis its 28 centromeres are visualised as 14 centromeric sites which again indicated pairing. The reduction in the number of centromeric sites prior to meiosis happens in the presence or absence of the *Ph1* locus.

At the onset of meiosis as the telomere bouquet forms in tetraploid and hexaploid wheat with or without *Ph1*, the centromeric sites reduce to seven visible centres (Martinez et al. 2003). The centromeres in the seven centres then undergo a chromatin remodeling process- the centromeric chromatin elongates (it appears to “decondense”). In hexaploid wheat, the seven centres resolve to 21 sites and in tetraploid wheat to 14 sites. To date, chromosome-specific markers in wheat do not exist to study the nature of these associations (either when the centromeres are in pairs or as seven centres). However data from studies of centromere behaviour in rye-wheat hybrids provides an insight. Do the centromeres pair randomly or non-randomly prior to meiosis in a rye-wheat hybrid which possesses 28 centromeres (21 wheat and seven rye) but no true homologues (Martinez-Perez et al. 2001; Prieto et al. 2004)? Either in the presence or absence of *Ph1* in the hybrid, the 21 wheat centromeres reduce to 7 centres in all tissues of the developing anther whether its meiocytes, tapetal or cell wall cells prior to meiosis. Thus the wheat centromeres are not pairing randomly prior to meiosis. Each wheat centromere must only be able to pair with two other wheat centromeres out of the potential 20 wheat centromeres whether *Ph1* is present or absent. The implication is that pairing of wheat centromeres prior to meiosis involves homoeologous centromeres. Thus in species like wheat and rice, chromosomes can be sorted both by their centromeres and telomeres. However in the absence of *Ph1*, the seven rye centromeres do not pair with these seven wheat centres while in the presence of *Ph1*, each of the seven wheat sites contains one rye centromere. Thus *Ph1* is specifically affecting rye centromere behaviour in relation to wheat centromere behaviour prior to meiosis (Prieto et al. 2004). Thus we now know that rye centromeres do not mirror the behaviour of wheat centromeres prior to meiosis and so can't be used as a marker for wheat centromere pairing (for example as described in later study of Corredor et al. 2007). Centromeres pair during replication (Jasencakova et al. 2001). Therefore since *Ph1* is affecting centromere pairing behaviour of rye centromeres, then by implication it is affecting the replication of rye centromeres with respect to wheat centromeres prior to meiosis. Thus observations on centromere pairing provide a further indication that *Ph1* is affecting the replication process.

At the subsequent telomere bouquet stage in the presence of *Ph1* in the rye-wheat hybrid, the centromeres associate as seven centres, with each centre

containing one of the seven rye centromeres. However in the absence of *Ph1*, the association of rye and wheat centromeres as seven centres occurs rarely as the centromeres undergo a remodelling process at this stage. As stated previously wheat chromosomes are being sorted both by their centromeres and telomeres. The advantage of sorting chromosomes via centromere pairing independently from the sites along rest of the chromosome may be that if crossover fails to occur between the paired chromosomes later in meiosis, it may allow chromosomes to still be segregated even through the chromosomes are not tethered by a crossover.

In the presence of *Ph1*, the centromeres in the seven centres then undergo the chromatin remodeling process and resolve as 28 unpaired sites. In the absence of *Ph1*, the chromatin remodeling process happens earlier as the rye centromeres should join the 7 wheat centromeres. In the absence of *Ph1*, centromeres resolve as 7 wheat pairs and 7 unpaired wheat and 7 unpaired rye centromeres. Thus after the chromatin remodeling stage, in the presence of *Ph1*, homoeologous centromeres are unable to pair with each other. These studies indicate the importance of the “remodeling” process in which the specificity of pairing at these sites changes after this stage in the presence and absence of *Ph1*. The chromatin remodeling of centromeres also seems to occur earlier in the absence of *Ph1* than its presence.

At the onset of meiosis, chromosomes start the process of condensing. However the subtelomeric constitutive heterochromatin in maize and rye which is already condensed, undergoes chromatin remodeling and elongates as the telomeres cluster to form a bouquet (Bass et al. 1997; Prieto et al. 2005). The constitutive heterochromatin at the centromere and subtelomeric regions undergoes similar conformation changes at a time when the rest of the chromosomes is initiating the condensation process. Recent data indicates that subtelomeric heterochromatin must undergo these conformational changes at the onset of meiosis in order for pairing and recombination to occur between chromosomes. In the presence of *Ph1*, the heterochromatin must be identical or near identical for the chromatin remodeling of the heterochromatin to occur on both homologues (Prieto et al. 2004; Colas et al. 2008). If the heterochromatin shows too much divergence then they are unable to remodel at the telomere bouquet stage and therefore the homologues are unable to pair and recombine via their telomere regions. The implication is that the homoeologous segments are unable to undergo remodeling in the presence of *Ph1* and therefore do not pair and recombine in contrast to the homologous segments. However, in the absence of *Ph1*, all related heterochromatin can remodel without the requirement of an identical or near identical heterochromatin on another chromosome at the telomere bouquet stage. Thus *Ph1* is diploidising the behaviour of heterochromatin. In the presence of *Ph1*, only near or identical heterochromatin can remodel. Since near or identical heterochromatin is found on homologues, only heterochromatin on homologues will be remodeled simultaneously. In the hybrids between wheat and related species, where there are no homologues, in the absence of *Ph1*, related heterochromatin is remodeled

synchronously at the telomere bouquet stage and the chromosomes pair, while in the presence of *Ph1*, heterochromatin does not undergo this synchronized remodeling at the telomere bouquet stage and the chromosomes don't pair. The presence of B chromosomes which can compensate for absence of *Ph1*, disrupt the synchronized replication of related heterochromatin (Pryor et al. 1980). Thus *Ph1*'s effect on heterochromatin links the process of chromatin remodeling with replication. In the absence of *Ph1*, heterochromatin remodeling can occur earlier both at the centromeres and sub-telomere regions than in its presence. Does *Ph1* affect the behaviour of heterochromatin along the rest of the chromosome arms? Studies reveal that the initiation of condensation is more coordinated in the presence than the absence of *Ph1* (Prieto et al. 2004). This effect on coordination of initiation of condensation may simply reflect a change in the coordination of an earlier process, namely premeiotic replication. Thus true homologues can initially be in different conformational states. This increases the chance of homoeologous pairing associations initially. At later stages, the condensation of the homologues is similar. The diploidisation of heterochromatin and the coordination of the initial phases of condensation explains why there are fewer associations between multiple chromosomes observed in Holm's synapsis studies in the presence of *Ph1* than its absence. In summary, the *Ph1* locus affects a process during replication as revealed by centromere behaviour, heterochromatin remodeling, coordination of the initiation of condensation (which may be linked also to replication) and synapsis and recombination during meiosis.

Recent molecular analyses have defined the *Ph1* locus to a region containing a cyclin-dependent kinase complex (*Cdk2-like* genes) related to *Cdk2* from humans and mouse which has been disrupted by the insertion of a segment of subtelomeric heterochromatin (Griffiths et al. 2006; Al-Kaff et al. 2007). *Cdk2* has been extensively studied because it controls chromatin remodeling at replication (Cohen et al. 2006). However more recently *Cdk2* has been shown to control the remodeling of heterochromatin by altering the binding of the heterochromatin protein (HP1) to such regions (Hale et al. 2006). This change in binding means that the heterochromatin "elongates". Disruption of *Cdk2* has also been shown to have a major effect on meiosis resulting in sterility and non-homologous synapsis. *Cdk2* has been shown to be required for the production of synaptonemal complex proteins presumably through its ability to phosphorylate factors involved in controlling their production (Ortega et al. 2003). More recently still, *Cdk2* has been shown to recruit mismatch repair proteins to double strand breaks during early stages of meiosis (Ward et al. 2007). These observations on *Cdk2* function are similar to that described above for the *Ph1* locus, namely in affecting replication, chromatin remodelling and synapsis/recombination.

Interestingly, it has been shown that the 5B *Cdk2-like* gene complex is suppressing expression of the corresponding *Cdk2-like* loci on 5A and 5D in bread wheat (Al-Kaff et al. 2007). The 5B *Cdk2-like* complex contains pseudo-genes which are being transcribed. The presence of multiple copies of *Cdk2-like*

genes including expressed pseudogenes may explain why it is difficult to generate mutant *Ph1* phenotypes with EMS treatments which only yield point mutations. Transcriptional analysis from the homoeologous genes reveals that most of it comes from the B genome genes. If these genes are deleted, then the loss of transcription is compensated by an increase in transcription from either or both on the genes in the 5A or 5D genomes. Deletion of the 5B *Cdk2-like* locus therefore results in activation of the 5A and 5D *Cdk2-like* loci. Thus the 5B locus is dominant with respect to the 5A and 5D loci. Feldman reported that chromosome 5A and 5D may carry homoeologous loci related to the *Ph1* locus which affect chromosome pairing (Feldman 1966). Deletion of 5D or multiple copies of the 5A chromosome could affect the regular pairing observed in wheat. The progenitor of the B genome of hexaploid and tetraploid wheat is likely to be related to *Ae. speltoides*. However its genome does not carry anything which can substitute or compensate for *Ph1* activity on 5B. As Feldman (1966) proposed, it is likely that *Ae. speltoides* and *T. urartu* produced a hybrid exhibiting homoeologous pairing. Therefore a polyploidisation event on 5B gave rise to the *Ph1* locus which suppresses homoeologous pairing activity.

Further dissection of the locus is required to assess whether all the *Cdk2-like* genes (seven of them) on 5B are involved in producing the *Ph1* phenotype, which of the 5B copies are involved in suppressing the expression of the 5A and 5D loci, whether the *Ph1* mutant phenotype involves a contribution of 5A and 5D loci or whether it simply reflects the loss of the 5B locus and finally the role of the segment of heterochromatin. However at present, it is difficult to envisage how to further dissect this region. Identification of more deletion breakpoints which fall within this specific region would require screening mutagenised populations of more than half a million individuals in size in order to generate deletion breakpoints distributed every 250 Kb. This also assumes that the distribution of breakpoints will be random. However the presence of the segment of heterochromatin may affect the occurrence of breakpoints in the flanking regions. There are a total of 14 *Cdk2-like* genes on 5B, 5A and 5D. It is also difficult to envisage mutagenising each *Cdk2-like* gene individually through EMS treatment to yield point mutations and then recombining different combinations of mutagenised genes. Equally exploiting RNAi based approach would knock out all expression of copies on 5A, 5B and 5D and could well lead to a gross phenotype similar to that observed in mice, namely gross disruption of meiosis and sterility rather than the subtle *Ph1* effect observed with deletion just the 5B locus. Preliminary data suggests that knocking out all *Cdk2-like* genes may indeed mirror the mice situation.

To date as stated, no chromosome substituted for 5B in hexaploid wheat from any wild species has been shown to compensate fully for the *Ph1* locus. The 5G chromosome from the wheat relative *T. timopheevi*, when substituted for 5B does provide a similar effect (Ozkan and Feldman 2001). *T. timopheevi* is likely to have arisen from wide crossing between the same progenitors which gave rise to tetraploid wheat 0.5 million years earlier. Preliminary studies suggest that the *Ph1* locus on 5B arose by duplication of *Cdk2-like* genes and

the insertion of the heterochromatin segment following polyploidisation (Al-Kaff et al. 2007). Moreover the *Ph1* type locus in *T. timopheevi* seems similar in structure to that present in bread wheat, therefore suggesting that the *Ph1* locus arose once (Griffiths et al. 2006).

8.7 Exploitation of Chromosome Pairing Loci

Are loci such as *Ph1* important for plant breeding? This topic has been extensively reviewed recently for example by Qi et al. (2007). Therefore this section will focus specifically on aspects of the *Ph1* developing story. Wheat has been generated carrying deletions encompassing the *Ph1* locus enabling pairing to occur between related chromosomes (Sears 1977; Roberts et al. 1999; Al-Kaff et al. 2007). Wheat lacking *Ph1* have been used in the last 25 years for breeding. Interspecific hybrids made between wheat lacking *Ph1* and its wild relatives often exhibit extensive pairing at metaphase I between the homoeologues. This observation is of practical importance for plant breeding as there is the potential to generate recombinant chromosomes composed of part wheat and part non-wheat chromatin, allowing the introgression of exotic genetic material from wild relatives. Sears transferred disease resistance loci from wild relatives of wheat by recombining the chromosomes with those of wheat in lines *Ph1* deficient (Sears 1972). More recently CIMMYT exploited *Ph1* to introgress from closely related species (William et al. 2005). Unfortunately an important complication is that there is still a hierarchy in preferred pairing partners even in the absence of *Ph1*. Most pairing observed at metaphase I (80%) in hexaploid wheat interspecific hybrids lacking *Ph1* locus occurs -between wheat chromosomes from the A and D genomes (Jauhar et al. 1991). A second problem in using wheat lacking *Ph1* for breeding is that -the chromosomes in such lines are heavily rearranged, making it difficult for recombination to occur between wheat chromosomes and related chromosomes. Hence the stocks are maintained in a heterozygous situation. Therefore the deployed of the *Ph1* deletion line in breeding is not trivial. If it is possible to design methods that selectively silence the whole 5B *Cdk2-like* complex rather than the 5A or 5D loci, then transgenic plants expressing dsRNA targeting the 5B locus could be generated. However the constitutive silencing of *Ph1* will still eventually lead to lines carrying rearranged chromosomes. Moreover the constitutive silencing of the 5B locus must also be accompanied by the activation of the 5A and 5D loci whose overexpression may be contributing to the *Ph1* mutant phenotype rather than a silencing of all *Cdk2-like* expression from the 5A, 5B and 5D loci. Ideally one would like to be able to switch off *Ph1* in hybrid plants between wheat and a wild relative and then switch it back on in the BC plants so that stable recombinant lines could be recovered. If *Ph1 Cdk2-like* genes do function similar to CDK2 then it should be possible to design drugs which could be delivered directly into the immature inflorescence of the hybrid plant, thereby inhibiting

Ph1 function only during meiosis. There are many compounds which are known to activate CDK2 in mammals and yeast, such chemicals could be delivered into the inflorescence and then scored for the ability to reproduce the *Ph1* mutant phenotype. This approach could have a major impact on breeding strategies as it may not just affect homoeologous recombination but also homologous recombination as well. The alternative approach is to exploit the suppressors of *Ph1* activity which have been described in the genome of *A. speltoides* (Dvorak et al. 2006). When the chromosomes carrying these loci are introduced into wheat, they induce homoeologous pairing even in the presence of the *Ph1* locus. The identification of genes responsible may allow them to be exploited to regulate *Ph1* expression and hence pairing and recombination in wheat.

References

- Akhunov, E., Goodyear, A., Geng, S., Qi, L.-L., Echalier, B., Gill, B., Miftahudin, P., Gustafson, P., Lazo, G., Chao, S., Anderson, O., Linkiewicz, M., Dubcovsky, J., Rota, M., Sorrells, M., Zhang, D., Nguyen, H., Kalavacharla, V., Hossain, K., Kianian, S., Peng, J., Lapitan, N., Gonzalez-Hernandez, J., Anderson, J., Choi, D.-W., Close, T., Dilibrigli, M., Gill, K., Walker-Simmons, K., Steber, C., McGuire, P., Qualset, C. and Dvorak, J. (2003) The organisation and rate of evolution of wheat genomes are correlated with recombination rates along the chromosome arms. *Genome Res.* 13, 753–763.
- Al-Kaff, N., Knight, E., Bertin, I., Foot, T., Hart, N., Griffiths, S. and Moore, G. (2007) Detailed dissection of the chromosomal region containing the *Ph1* locus in wheat *Triticum aestivum* L.: with deletion mutants and expression profiling. *Ann. Bot.* doi:10.1093/aob/mcm252.
- Bass, H.W., Marshall, W.F., Sedat, J.W., Agard, D.A. and Cande, W.Z. (1997) Telomere cluster de novo before the initiation of synapsis; a three-dimensional spatial analysis of telomere positions before and during meiotic prophase. *J. Cell Biol.* 137, 5–18.
- Boden, S., Shadiac, N., Tucker, E., Langridge, P. and Able, J. (2007) Expression and functional analysis of TaAsyl during meiosis of bread wheat *Triticum aestivum*. *BMC Mol. Biol.* 8, 65.
- Bowers, J.E., Chapman, B.A., Rong, J. and Paterson, A.H. (2003) Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. *Nature* 422, 433–438.
- Chikashige, Y., Ding, D.Q., Imai, Y., Yamamoto, M., Haraguchi, T. and Hiraoka, Y. (1997) Meiotic nuclear reorganization switching the position of centromeres and telomeres in the fission yeast *Scizaccharomyces pome*. *EMBO J.* 16, 193–202.
- Cohen, P.E., Pollack, S.E. and Pollard, J.W. (2006) Genetic analysis of chromosome pairing, recombination and cell cycle control during first meiotic prophase in mammals. *Endocrine Reviews* 27, 398–426.
- Colas, I., Shaw, P., Prieto, P., Wanous, M., Spielmeier, W., Mago, R. and Moore, G. (2008) Effective chromosome pairing requires chromatin remodelling at the onset of meiosis. *Proc. Natl. Acad. Sci. USA* 105, 6075–6080.
- Corredor, E., Lukaszewski, A., Pachon, P., Allen, D.C. and Naranjo, T. (2007) Terminal regions of wheat chromosomes select their pairing partners in meiosis. *Genetics* 177, 699–706.
- Crismani, W., Baumann, U., Sutton, T., Langridge, P. and Able, J. (2006) Microarray expression analysis of meiosis and microsporogenesis in hexaploid bread wheat. *BMC Genomics* 7, 267.

- Curtis, C.A., Lukaszewski, A.J. and Chrzastek, M. (1991) Metaphase I pairing of deficient chromosomes and genetic mapping of deficiency breakpoints in common wheat. *Genome* 34, 553–560.
- Dong, C., Whitford, R. and Langridge, P. (2002) A DNA mismatch repair gene links to the *Ph2* locus in wheat. *Genome* 45, 116–124.
- Dover, G.A. and Riley, R. (1972) Prevention of pairing of homeologous meiotic chromosomes of wheat by an activity of supernumerary chromosomes of *Aegilops*. *Nature* 240, 159–161.
- Dubcovsky, J., Luo, M.-C. and Dvorak, J. (1995) Differentiation between homoeologous chromosomes 1A of wheat and 1Am of *Triticum monococcum* and its recognition by the wheat *Ph1* locus. *Proc. Natl. Acad. Sci. USA* 92, 6645–6649.
- Dvorak, J., Deal, K.R. and Luo, M.C. (2006) Discovery and mapping of wheat *Ph1* suppressor. *Genetics* 174, 17–27.
- Feldman, M. (1966) The effect of chromosomes 5B, 5D, and 5A on chromosomal pairing in *Triticum aestivum*. *Proc. Natl. Acad. Sci. USA* 55, 1447–1453.
- Feldman, M. (1993) Mode of action of *Ph1* in wheat. *Crop Sci.* 33, 894–897.
- Fu, H. and Dooner, H. (2002) Intrasppecific violation of genetic colinearity and its implications in maize. *Proc. Natl. Acad. Sci. USA* 99, 9573–9578.
- Griffiths, S., Sharp, R., Foote, T., Bertin, I., Wanous, M., Reader, S., Colas, I. and Moore, G. (2006) Molecular characterisation of *Ph1* as a major chromosome pairing locus in polyploid wheat. *Nature* 439, 749–752.
- Hale, T.K., Contreras, L.A. Morrison, A. and Herrera, R.E. (2006) Phosphorylation of the linker Histone H1 by Cdk regulates its binding to HP1. *Mol. Cell.* 22, 693–699.
- Holm, P.B. (1986) Chromosome pairing and chiasma formation in allohexaploid wheat, *Triticum aestivum*, analysed by spreading of meiotic nuclei. *Carlsberg Res. Commun.* 51, 239–294.
- Holm, P.B. (1988) Chromosome pairing and synaptonemal complex formation in hexaploid wheat nullisomic for chromosome 5B. *Carlsberg Res. Commun.* 53, 91–110.
- Holm, P.B. and Wang, X. (1988) The effect of chromosome 5B on synapsis and chiasma formation in wheat, *Triticum aestivum* cv Chinese Spring. *Carlsberg Res. Commun.* 53, 191–208.
- Jasencakova, Z., Meister, A. and Schubert, I. (2001) Chromatin organization and its relation to replication and histone acetylation during the cell cycle in barley. *Chromosoma* 110, 83–92.
- Jauhar, P.P., Rieva-Lizarazu, O., Dewey, W.G., Gill, B.S., Crane, C.F. and Bennett, J.H. (1991) Chromosome pairing relationships among the A, B and D genomes of bread wheat. *Theor. Appl. Genet.* 82, 441–449.
- Jones, G.H. (1984) The control of chiasma distribution. *Symp. Soc. Exp. Biol.* 38, 293–320.
- King, J., Armstead, A., Donnison, A., Roberts, L., Harper, J., Skot, K., Elborough, K. and King, I. (2007) Comparative analyses between *Lolium/Festuca* introgression lines and rice reveal the major fraction of functionally annotated gene models are located in recombination poor/very poor regions of the genome. *Genetics* 177, 547–606.
- Korzun, I. and Kunzel, G. (1996) The physical relationship of barley chromosome 5 (1H) to the linkage group of rice chromosome 5 and 10. *Mol. Gen. Genet.* 252, 225–231.
- Kopecky, D., Allen, D.C., Duchoslav, M., Dolezel, J. and Lukaszewski, A.J. (2007) Condensation of rye chromatin in somatic interphase nuclei of *Ph1* and *ph1b* wheat. *Cytogenet Genome Res.* 119, 263–267.
- Kunzel, G., Korzun, L. and Meister, A. (2000) Cytogenetically integrated physical restriction length polymorphism maps of the barley genome based on translocation breakpoints. *Genetics* 154, 397–412.
- Lloyd, A.H., Milligan, A.S. Langridge, P. and Able, A. (2007) TaMSH7:A cereal mismatch repair gene that affects fertility in transgenic barley (*Hordeum vulgare*). 2007 BMC Plant Biology doi:10.1186/1471-2229-7-67.

- Luo, M.-C., Dubcovsky, J. and Dvorak, J. (1996) Recognition of homoeology by the wheat *Ph1* locus. *Genetics* 144, 1195–1203.
- Lukaszewski, A.J. (1997) The development and meiotic behaviour of asymmetrical isochromosomes in wheat. *Genetics* 145, 1155–1203.
- Maestra, B., de Jong, J.H., Shepherd, K. and Naranjo, T. (2002) Chromosome arrangement and behaviour of rye homologous telosomes at the onset of meiosis in disomic wheat 5RL addition lines with and without the *Ph1* locus. *Chromosome Res.* 10, 655–667.
- Martinez, M., Cuñado, N., Carcelén, N. and Romero, C. (2001) The *Ph1* and *Ph2* loci play different roles in the specific behaviour of hexaploid wheat *Triticum aestivum*. *Theor. Appl. Genet.* 102, 751–758.
- Martinez-Perez, E., Shaw, P., Reader, S., Aragon-Alcaide, L., Miller, T. and Moore, G. (1999) Homologous chromosome pairing in wheat. *J. Cell Sci.* 112, 1761–1769.
- Martinez-Perez, E., Shaw, P. and Moore, G. (2000) Polyploidy induces centromere association. *J. Cell Biol.* 148, 233–238.
- Martinez-Perez, E., Shaw, P. and Moore, G. (2001) The *Ph1* locus is needed to ensure specific somatic and meiotic centromere association. *Nature* 411, 204–207.
- Martinez-Perez, E., Shaw, P., Aragon, L. and Moore, G. (2003) Chromosomes form into seven groups in hexaploid and tetraploid wheat as a prelude to meiosis. *Plant J.* 36, 21–29.
- Materson, J. (2003) Stomatal size in fossil plants: Evidence of polyploidy in majority of Angiosperms. *Science* 264, 421–424.
- Mikhailova, E.I., Naranjo, T., Shepherd, L., Wennekes-van Eden, J., Heyting, C. and de Jong, J.H. (1998) The effect of the wheat *Ph1* locus on chromosome organization and meiotic chromosome pairing analysed by genome painting. *Chromosoma* 107, 339–350.
- Mikhailova, E.I., Sosnikhina, S.P., Kirillova, G., Tikhoniz, O.A., Smirnov, V.G., Jones, R.N. and Jenkins, G. (2001) Nuclear dispositions of subtelomeric and pericentromeric chromosomal domains during meiosis in asynaptic mutants of rye (*Secale cereale* L.). *J. Cell Sci.* 114, 1875–1882.
- Moore, G. (1995) Cereal genome evolution: pastoral pursuits with lego genomes. *Curr Opin. Genet. Dev.* 5, 717–724.
- Ortega, S., Prieto, I., Odajima, J., Martin, A., Dubus, P., Sotillo, R., Barbero, J., Malumbres, M. and Barbacid, M. (2003) Cyclin-dependent kinase 2 is essential for meiosis but not for mitotic cell division in mice. *Nat. Genet.* 35, 25–31.
- Ozkan, H. and Feldman, M. (2001) Genotypic variation in tetraploid wheat affecting homologous pairing in hybrids with *Aegilops peregrine*. *Genome* 44, 1000–1006.
- Niwa, O., Shimanuki, M. and Mika, F. (2000) Telomere-led bouquet formation facilitates homologous chromosome pairing and restricts ectopic interaction in fission yeast meiosis. *EMBO J.* 19, 3831–3840.
- Page, S.L. and Hawley, R.S. (2003) Chromosome choreography: the meiotic ballet. *Science* 301, 785–789.
- Paux, E., Roger, D., Badaeva, E., Gay, G., Bernard, M., Sourdille, P. and Feuillet, C. (2006) Characterising the composition and evolution of homoeologous genomes in hexaploid wheat through BAC end sequencing of chromosome 3B. *Plant J.* 48, 463–474.
- Prieto, P., Shaw, P. and Moore, G. (2004) Homologue recognition during meiosis is associated with change in chromatin conformation. *Nat. Cell Biol.* 6, 906–908.
- Prieto, P., Santos, A.P., Moore, G. and Shaw, P. (2004b) Chromosomes associate premeiotically and in xylem vessel cells via their telomeres and centromeres in diploid rice (*Oryza sativa*). *Chromosoma* 112, 300–307.
- Prieto, P., Moore, G. and Reader, S. (2005) Control of conformation changes associated with homologue recognition during meiosis. *Theor. Appl. Genet.* 111, 505–510.
- Pryor, A., Faulkner, K., Rhoades, M.M. and Peacock, W.J. (1980) Asynchronous replication of heterochromatin in maize. *Proc. Natl. Acad. Sci. USA* 77, 6705–6709.
- Qi, L., Friebe, B., Zhang, P. and Gill, B.S. (2007) Homoeologous recombination, chromosome engineering and crop improvement. *Chromosome Res.* 15, 3–19.

- Riley, R. and Chapman, V. (1958) Genetic control of the cytological diploid behaviour of hexaploid wheat. *Nature* 182, 712–715.
- Roberts, M., Reader, S., Dalglish, C., Miller, T., Foote, T., Fish, L., Snape, J. and Moore, G. (1999) Induction and characterisation of *Ph1* wheat mutants. *Genetics* 199, 153, 1909–1918.
- Sanchez-Moran, E., Benavonte, E. and Orellana, J. (2001) Analysis of karyotypic stability of homoeologous-pairing mutants in allopolyploid wheat. *Chromosoma* 110, 371–377.
- Scherrer, B., Isidore, E., Klein, P., Kim, J.-S., Bellec, A., Chalhou, B., Keller, B. and Feuillet, C. (2005) Large intraspecific haplotype variability at the *Rph7* locus from rapid and recent divergence in the barley genome. *Plant Cell* 17, 361–374.
- Sears, E. (1972) Agropyron-wheat transfers through induced homoeologous pairing. *Can. J. Genet. Cytol.* 14, 746.
- Sears, E. (1977) An induced mutant with homoeologous pairing in common wheat. *Can. J. Genet. Cytol.* 19, 585–593.
- Sutton, T., Whitford, R., Baurmann, U., Dong, M., Able, J. and Langridge, P. (2003) The *Ph2* pairing homoeologous locus in wheat (*Triticum aestivum*); identification of candidate meiotic genes using a comparative genetics approach. *Plant J.* 36, 443–456.
- Tresses-Sticken, E., Dresser, M.E. and Scherthan, H. (1999) Bouquet formation in budding yeast: initiation of recombination is not required for meiotic telomere clustering. *J. Cell Sci.* 112, 651–658.
- Ward, J.O., Reinholdt, L.G., Motley, W.W., Niswander, L.M., Deacon, D.C., Griffin, L.B., Langlais, K.K., Backus, V.L., Schimenti, K.J., O'Brien, M.J., Eppig, J.J. and Schimenti, J.C. (2007) Mutation in mouse *He10*, an E3 Ubiquitin ligase disrupts meiotic crossing over. *PLoS Genetics* 3:e139 doi:10.1371/journal.pgen.003011139.
- William, H.M., Singh, R.P., Trethowan, R., Van Ginkel, M., Pellegrinshi, A., Huerta-Espin, A. and Hosington, D. (2005) Biotechnology applications for wheat improvement CIM-MYT. *Turk J. Agri.* 29, 113–119.
- Zickler, D. and Kleckner, N. (1999) Meiotic chromosomes: integrating structures. *Ann. Rev. Genet.* 33, 603–754.

Chapter 9

A Toolbox for Triticeae Genomics

Etienne Paux and Pierre Sourdille

Abstract In the last two decades, progress in cereal genomics has been remarkable, enabling a better understanding of the structure and function of the cereal genomes. However, significant advances mainly concerned rice and maize, whereas for the Triticeae species, namely wheat, barley and rye, the development of genomic tools and resources has long been hampered by the size and the complexity of their genomes. Recently, new technologies have allowed the development of a toolbox for Triticeae genomics comparable to what is available for rice and maize. Triticeae scientists and breeders can now benefit from a wide range of tools, including molecular markers, genetic maps, EST sequences, microarrays, BAC libraries and transformation systems that can be applied to structural, functional, evolutionary and comparative genomic studies of the Triticeae genomes.

9.1 Introduction

Genomics of the Triticeae species, namely wheat, barley and rye, has been hampered for a long time by the size (17 Gb for wheat, 4.9 Gb for barley and 7.6 Gb for rye; Zonneveld et al. 2005) and the complexity (>80% of repetitive sequences, polyploidy in wheat species) of their genomes. Recently, large international efforts in genomics and high throughput technologies have allowed the development of molecular markers and genetic maps, as well as physical maps, large scale sequencing, and microarrays in the Triticeae. These tools are now widely used for gene mapping and cloning, diversity and evolutionary studies, genome structure exploration or marker-assisted selection.

Cereal genomics has been the subject of several reviews (Appels et al. 2003; Gupta and Varshney 2004; Henry 2007; Varshney et al. 2006a) that mainly

E. Paux (✉)

INRA UMR1095 Génétique, Diversité and Ecophysiologie des Céréales, Domaine de Crouël, 234 Avenue du Brézet, 63100 Clermont-Ferrand France
e-mail: etienne.paux@clermont.inra.fr

focused on rice and maize while the specific field of Triticeae genomics has not been extensively reviewed so far. In this chapter, we provide an overview of some of the most important genomics tools available for the Triticeae with a first focus on the description of different types of molecular markers followed by a brief review of Expressed Sequenced Tags (ESTs) and their application for expression studies or mapping and finally an overview on the resources available for physical mapping.

9.2 Molecular Markers

If initial marker development focused on RFLP (Restriction Fragment Length Polymorphism) markers, their numerous drawbacks have rapidly lead to the burst of a new generation of PCR- or array-based markers such as microsatellites (Simple Sequence Repeats, SSRs), AFLPs (Amplification Fragment Length Polymorphisms) and more recently DArTs (Diversity Array Technology), repeat-based markers and SNPs (Single Nucleotide Polymorphisms). The next paragraphs provide a description of the technical basis of the different techniques and examples of their application for genetics and genomics studies in the Triticeae. A summary of the features and number of each type of markers is provided at the end of this paragraph (Table 9.3)

9.2.1 *Restriction Fragment Length Polymorphism (RFLP) Clones*

Restriction Fragment Length Polymorphism (RFLP) is a technique that uses restriction endonucleases (RE) to cut DNA at specific (generally 4–6 bp) recognition sites. The resulting fragments are separated according to their molecular size using agarose gel electrophoresis and after transfer to a nylon membrane (Southern blotting) the DNA is hybridized with labeled low-copy DNA fragments used as probes. The presence, absence or size modification of fragments resulting from polymorphism in RE recognition sites are used to compare species, varieties or segregating populations. Initially used to identify mutations in thermo-sensitive adenovirus strains (Grodzicker et al. 1974), RFLPs have largely advanced due to the development of the human genetic map (Botstein et al. 1980). Since then, the technique was extended to most of the crop plants including maize (Burr et al. 1983), tomato (Tanksley et al. 1988) and soybean (Keim et al. 1990).

In the Triticeae, RFLP markers were primarily used to build genetic maps (Cadalen et al. 1997; Chao et al. 1989; Heun et al. 1991) for gene and QTL detection. Despite the fact that these markers cannot be automated and can thus only be used on samples of limited size (few hundreds), they remain essential for comparative mapping studies since DNA probes from one species can be efficiently used on related species (Devos et al. 1993). Gene probes

derived from cDNA libraries are particularly useful for comparisons between related species and they allowed the first colinearity studies between chromosomes of several grass species using rice as a pivotal genome (Moore et al. 1995).

Several genomic as well as cDNA libraries were developed as source of probes for RFLP analysis in Triticeae leading to several thousands of publicly available probes (see Graingenes database: www.wheat.pw.usda.gov and Table 9.1). Genomic clones were mainly developed from methylation sensitive restriction endonucleases, such as *Pst*I and *Hpa*II, to increase the representation of low copy sequences (Cheung et al. 1992; Devey and Hart 1993; Devos and Gale 1993; Gill et al. 1991; Heun et al. 1991; Jahoor et al. 1991; Liu and Tsunewaki 1991). cDNA clones were obtained from various tissues of barley (Heun et al. 1991; Kleinhofs et al. 1993), oat (Heun et al. 1998; Sharp et al. 1989) and wheat (Devos et al. 1993; Sharp et al. 1989) and have been used in a number of studies as well as the basis for EST marker development (see Section 9.3).

With RFLPs, a positive correlation is found between the size of the fragments and the polymorphism frequency generated by different enzymes suggesting that polymorphism is mainly due to chromosomal rearrangements such

Table 9.1 List of publicly available DNA clones that can be used to detect RFLPs in Triticeae species

Type of marker	Species	Type of DNA	Nomenclature	# of markers	References
RFLP	<i>H. vulgare</i>	cDNA	ABC	419	Kleinhofs et al. 1993
	<i>H. vulgare</i>	gDNA	ABG	155	Kleinhofs et al. 1993
	<i>H. vulgare</i>	cDNA	BCD	977	Heun et al. 1991
	<i>A. sativa</i>	cDNA	CDO	538	Heun et al. 1991
	<i>T. aestivum</i>	gDNA	FBA, FBB	300	Nelson et al. 1995a, b, c
	<i>T. aestivum</i>	gDNA	GBX	134	Mingeot & Jacquemin 1999
	<i>T. aestivum</i>	gDNA	GLK	457	Liu & Tsunewaki 1991
	<i>Ae. tauschii</i>	gDNA	KSU	223	Gill et al. 1991
	<i>T. aestivum</i> , <i>T. durum</i>	cDNA	MTA, MTD	16	P Joudrier unpublished
	<i>H. vulgare</i>	gDNA	MWG	405	Graner et al. 1991
	<i>T. aestivum</i>	cDNA, gDNA	PSR	350	Gale et al. 1995
	<i>T. aestivum</i>	gDNA	TAM	73	Devey & Hart 1993
	<i>T. aestivum</i>	gDNA	WG	81	Heun et al. 1991
Total				4128	

as insertions or deletions (Graner et al. 1991, 1990; Kammorgan et al. 1989). Some restriction enzymes such as *DraI*, *EcoRI*, *EcoRV* and *HindIII* are better to detect these changes than others. In contrast to species like maize, polymorphism is low in the Triticeae (Cadalen et al. 1997; Chao et al. 1989; Graner et al. 1990) ranging from 9% to 50%. This often leads to perform genetic mapping using crosses with a wide genetic basis, inter-specific crosses or crosses involving synthetic wheats that show higher levels of polymorphism (about 70%; Marino et al. 1996; Nelson et al. 1995a,b,c). Most of the RFLP markers are able to detect the three homoeologous copies when working on hexaploid bread wheat and the level of polymorphism varies between the three genomes, the B genome being the most polymorphic followed by the A and the D genome (Cadalen et al. 1997; Chao et al. 1989). Strikingly, some chromosomes exhibit very few polymorphisms like chromosome 4D of bread wheat.

The high cost of genotyping with RFLPs together with their low throughput capacity and low polymorphism detection power resulted in a very limited use of these markers by plant breeders and promoted the development of new types of polymerase chain reaction (PCR)-based molecular markers. RFLP markers are nowadays mainly used for comparative genomics and map-based cloning.

9.2.2 Simple Sequence Repeat (SSR) Markers

Microsatellites, also referred to as Simple Sequence Repeats (SSRs), Variable Number of Tandem Repeats (VNTRs) or Variable Simple Sequence Motifs (VSSMs), were first described in eukaryotes in the early eighties (Hamada et al. 1982). They are defined as stretches of DNA that consist of only one or a few (maximum of six) tandemly repeated nucleotides, such as poly(A/T) or poly(GT/CA) of less than 100 bp in length with a minimal length of 12 bp (Tautz 1993; Tautz and Renz 1984). These types of simple sequences were shown to be repetitive and widely distributed in many eukaryotic genomes (Tautz et al. 1986; Dover and Tautz 1986; Greaves and Patient 1985; Hamada et al. 1982; Tautz and Renz 1984). SSRs were reported to be involved in a range of functions including gene regulation (Hentschel 1982; Wang et al. 1979; Weintraub and Groudine 1976), signals for gene conversion and recombination (Goodman 1985; Jeffreys et al. 1985), and the replication of telomeres (Blackburn 1984).

Microsatellites are divided into different categories according to their composition (Weber 1990): perfect, imperfect or compound (two or more different linked SSRs) repeats. They are also classified into two classes based on their origins. Some are developed from enriched genomic DNA libraries (Edwards et al. 1996; Ostrander et al. 1992) and are thus named genomic SSRs (g-SSRs). Most of them have neither genic function nor close linkage to coding regions (Metzgar et al. 2000), and their developing process is very tedious and expensive. The second class of SSRs is developed from the expressed regions of the

genome and is named EST-SSRs. In general, EST-SSR markers produce high quality patterns, but give a lower level of polymorphism compared to that of genomic SSRs (Holton et al. 2002; Thiel et al. 2003). An important feature of the EST-SSR markers is that they can be rapidly developed from EST databases at low cost. In addition, because they originate from expressed regions, they can be used for selecting genes of interest through marker assisted selection (MAS). Moreover, since they are developed from single copy, coding regions, they can be used, like RFLPs, for comparative and phylogenetic studies (Zhang et al. 2005, 2006, 2007).

In the past ten years, a large number of SSR markers have been developed for barley (For a review see Hearnden et al. 2007), rye (Saal and Wricke 1999), and wheat (Gupta et al. 2003; Gupta and Varshney 2000; Guyomarc'h et al. 2002a,b; Röder et al. 1998; Varshney et al. 2002, see Table 9.2). The most common SSRs identified in the different studies, were GA and GT dinucleotide repeats. The density of GA (38–59%) and GT (20–34%) in these different species ranged from one SSR every 212 kb to one per 704 kb (Varshney et al. 2002). In bread wheat, estimates of the SSR frequency per haploid genome were one SSR each 3.6×10^4 bp for (GA)_n and each 2.3×10^4 bp for (GT)_n (Röder et al. 1995). In barley, the frequency of EST-SSRs ranged from one every 3.4 kb (Kantety et al. 2002) to one SSR every 7.5 kb (Kantety et al. 2002; Varshney et al. 2002) while in wheat, estimates ranged from one EST-SSR every 1.5 kb (Morgante et al. 2002) to one every 17.2 kb (Gao et al. 2003). Overall, an average of one microsatellite every 6–7 kb seems to be a good estimate for SSR frequency in the Triticeae genomes.

Numerous studies demonstrated that g-SSRs show a high level of polymorphism compared to other types of molecular markers, especially RFLPs. In barley, it was shown that even if SSRs exhibit a lower diversity index compared to AFLP (0.521 and 0.937 respectively), they reveal a higher number of alleles at each locus making them very effective to study genetic relationships (Russell et al. 1997). In cultivated spelt wheat, Bertin et al. (2001) found that the mean PIC value was 0.64 while in common wheat, it was 0.71. Even if lower PIC value (0.57) are also reported for common wheat (Stachel et al. 2000), this is twice higher than what is revealed by RFLPs (Cadalen et al. 1997; Chao et al. 1989). Assessments of polymorphism of EST-SSRs were carried out in barley (Thiel et al. 2003), rye (Hackauf and Wehling 2002) and wheat (Eujayl et al. 2002; Gupta et al. 2003; Holton et al. 2002; Leigh et al. 2003; Nicot et al. 2004) showing that EST-SSRs give higher quality profiles compared to the g-SSRs but that the polymorphism level ranged from 7% to 55% which was lower compared to g-SSRs.

Despite the fact that SSRs are expensive to develop because extensive sequencing is needed at the beginning of a project, they are currently the most widely used type of markers in Triticeae for genetic mapping as well as for diversity analyses. During the last ten years, nearly four thousand SSRs have been developed for Triticeae species and have been released in the public domain (<http://wheat.pw.usda.gov/GG2/index.shtml>; Table 9.2) making them useful for the whole community and widely used in all genomics projects.

Table 9.2 List of publicly available SSRs for Triticeae species

Type of marker	Species	Type of DNA	Nomenclature	# of markers	References
SSRs	<i>T. aestivum</i>	gDNA	BARC	544	Cregan & Song unpublished
	<i>H. vulgare</i>	gDNA	Bmac	55	Ramsay et al. 2000
	<i>H. vulgare</i>	gDNA	Bmag	104	Ramsay et al. 2000
	<i>T. monococcum</i>	gDNA	CFA	55	Guyomarc'h et al. 2000
	<i>T. aestivum</i>	gDNA	CFB	98	Paux et al. 2006
	<i>Ae. tauschii</i>	gDNA	CFD	130	Guyomarc'h et al. 2000
	<i>T. aestivum</i>	cDNA	CFE	302	Zhang et al. 2005
	<i>T. aestivum</i>	cDNA	CWEM	48	Gao et al. 2003
	<i>T. aestivum</i>	cDNA	CWM	475	Gao et al. 2003
	<i>H. vulgare</i>	gDNA	EBmac	64	Ramsay et al. 2000
	<i>H. vulgare</i>	cDNA	GBM	496	Thiel et al. 2003, Varshney et al. 2006
	<i>Ae. tauschii</i>	gDNA	GDM	65	Pestsova et al. 2000
	<i>H. vulgare</i>	gDNA	GMS	15	Struss & Plieske 1998
	<i>T. aestivum</i>	cDNA, gDNA	GPW	113	Nicot et al. 2004
	<i>H. vulgare</i>	cDNA, gDNA	Hv	162	Becker & Heun 1995, Liu et al. 1996, Pillen et al. 2000
	<i>T. aestivum</i>	gDNA	PSP	44	Bryan et al. 1997
	<i>S. cereale</i>	gDNA	SCM	27	Saal & Wricke 1999
	<i>H. vulgare</i>	cDNA	UMB	76	Beaubien and Smith 2006
	<i>T. aestivum</i>	gDNA	WMC	834	Somers et al. 2004
	<i>T. aestivum</i>	gDNA	WMS	238	Röder et al. 1998
Total				3945	

9.2.3 Amplified Fragment Length Polymorphism (AFLP) Markers

Amplified Fragment Length Polymorphism is a technique that combines the use of two restriction endonucleases and the amplification of selected restricted fragments through PCR (Zabeau and Vos 1993; Vos et al. 1995). Briefly, it involves three steps: (1) DNA is first restricted and oligonucleotide adapters are ligated at both ends. *Mse*I, a four base pair cutter and *Eco*RI, a six base pair cutter are the two most frequently used enzymes; (2) restricted fragments are

then selectively amplified using the adapter and restriction site sequence as target sites for primer annealing. The selective amplification is achieved by the use of primers that extend into the restriction fragments from one to three random bases. Only the fragments in which the primer extensions match the nucleotides flanking the restriction sites can be amplified; (3) the amplified fragments are separated and analyzed on acrylamide gels. The method allows the specific co-amplification of a high number of restriction fragments (between 50 and 100) whatever the origin or the complexity of the DNA and without prior knowledge of the sequence.

AFLP fingerprinting has been widely used in Triticeae species. It has been exploited and proven to be powerful to assess genetic diversity in wheat (Barrett and Kidwell 1998; Barrett et al. 1998; Burkhamer et al. 1998) and barley (Ellis et al. 1997; Schut et al. 1997) and has also been extensively used to construct high-density genetic maps of barley (Becker et al. 1995; Qi and Lindhout 1997) and wheat (Quarrie et al. 2005) as well as to analyse quantitative traits in both species (Pakniyat et al. 1997; Powell et al. 1997; Quarrie et al. 2005). Despite the ease with which they may be developed AFLPs suffer from some drawbacks. For example, the level of detectable polymorphism is relatively low and few markers are codominant. In addition, preliminary observations in barley suggested that AFLPs generated after digestion of the DNA with the recommended enzyme pair (*EcoRI/MseI*) are frequently clustered in centromeric regions and at the edge of genomic segments that are already heavily represented by RFLPs (Becker et al. 1995). Moreover, despite their interest in genetic mapping and linkage analyses, AFLPs have been only rarely used for physical mapping in the Triticeae because sequence specific primers are difficult to develop from AFLP markers (Shan et al. 1999).

9.2.4 Repeat-Based Markers

Repeated sequences account for 5 to more than 80% of all eukaryotic genomes. They are particularly abundant in the Triticeae genomes (Flavell et al. 1977) where they represent about 80% (Li et al. 2004; Paux et al. 2006). This feature has long hampered positional cloning of genes as well as high throughput sequencing projects in barley and wheat. However, recent sequencing in the Triticeae provided an opportunity to examine closely the organization of genes and repeated sequences on large stretches of chromosomes (Panstruga et al. 1998; Rostoks et al. 2002; Taketa et al. 2000; Li et al. 2004; Sabot et al. 2005; Paux et al. 2006). This revealed that Transposable Elements (TEs) are nested within each other, thus providing a way to estimate the date of invasion of LTR-retrotransposons and providing a tool for evolution studies. Using such estimation methods, it was found that the *Mla* resistance locus in barley was invaded by BARE-1 elements over the last two million years (Wei et al. 2002). TE insertions can occur or TE can be deleted during the breeding process because

of recombination which can lead to polymorphism in the sequence between different varieties.

Transposable elements were always considered as a major limitation for molecular marker development. However, because they are widespread throughout the Triticeae genomes (Muniz et al. 2001; Suoniemi et al. 1996), they certainly represent a great potential for whole genome scanning and mapping. In addition, they are less subjected to selection pressure than genes and can thus tolerate higher levels of mutation including insertions/deletions which make them highly polymorphic, even in closely related cultivars.

Most of the methods using repeated sequences as a source of molecular markers are based on PCR with primers designed in conserved regions of the TEs. The first TE based methods were S-SAP (sequence-specific amplified polymorphism; Waugh et al. 1997), I-RAP (inter-retrotransposon amplified polymorphism; Kalendar et al. 1999) or REMAP (Retrotransposon-Microsatellite Amplified Polymorphism; Kalendar et al. 1999) (Fig. 9.1).

The S-SAP technique relies on a combination of AFLP and retrotransposon sequence-specific PCR (Fig. 9.1A). The method produces amplified fragments containing a retrotransposon long terminal repeat (LTR) sequence at one end and a flanking host restriction site at the other. The basic requirement is the proximity of a given sequence (conserved or not) to a relatively rare cutting restriction enzyme site to which adaptors can be ligated. By adjusting both the enzyme and/or number of selective nucleotides on the adaptor-homologous oligonucleotide and/or by incorporating an initial asymmetric PCR with a chosen primer, it is possible to amplify fragments from almost any oligonucleotide sequence. In S-SAP, high copy number and dispersion throughout the genome are advantages, providing more information per assay. The level of polymorphism detected is higher than for AFLPs, which is valuable for species such as wheat where polymorphism is low.

Both IRAP and REMAP examine polymorphism in retrotransposon insertion sites; IRAP between retrotransposons and REMAP between retrotransposons and microsatellites (Kalendar et al. 1999). The IRAP markers are generated by the proximity of two LTRs using outward-facing primers annealing to LTR target sequences (Fig. 9.1B). Retrotransposons may integrate in principle in either orientation into the genome, and hence any two members of a retrotransposon family may be found head-to-head, tail-to-tail, or head-to-tail. For the first two orientations, a single primer suffices to generate PCR products from elements sufficiently close to one another (Fig. 9.1B). To amplify intervening genomic DNA for elements in head-to-tail orientation, both 5' and 3' LTR primers must be used, resulting in the production of additional fragments originating from the LTR found in the other orientations as well.

In REMAP, markers result from the amplification of DNA segments located between LTRs proximal to simple sequence repeats (SSRs) such as microsatellites. The method utilizes one outward-facing LTR primer and a second primer designed in a microsatellite (Kalendar et al. 1999; Fig. 9.1C). Primers are designed and anchored to the microsatellites 3' terminus by the addition of a

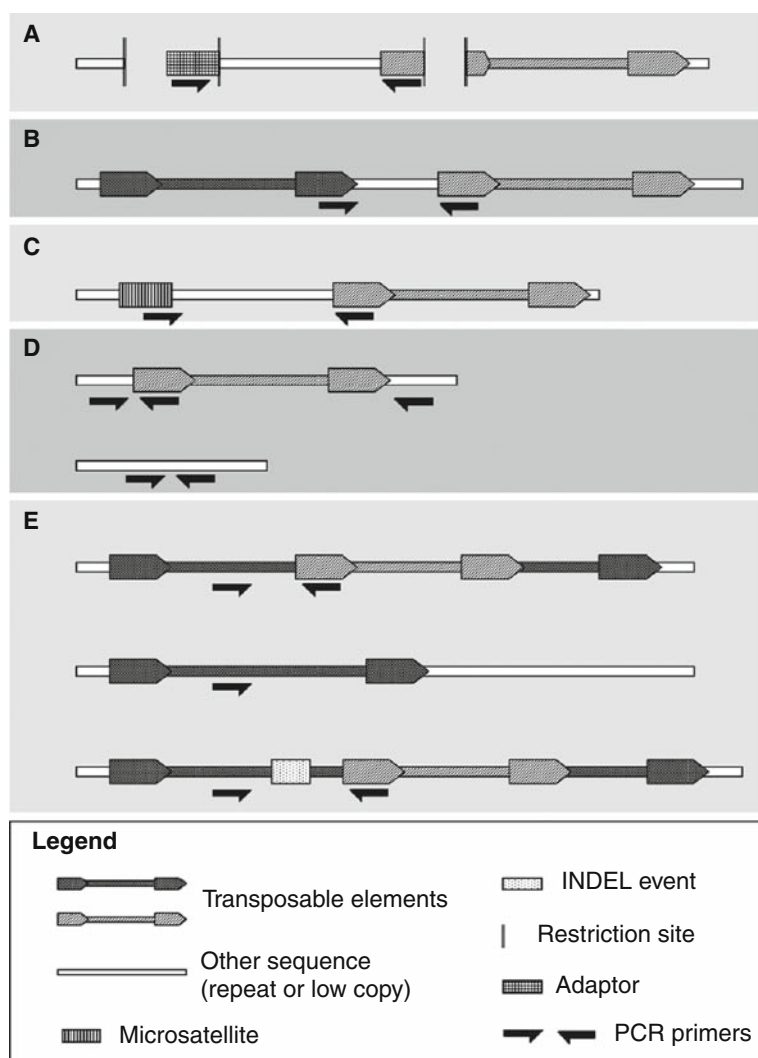


Fig. 9.1 Different types of repeat-based markers available for Triticeae: (A) S-SAP; (B) IRAP; (C) REMAP; (D) RBIP; (E) ISBP

single selective base at the 3' end. This was done to avoid polymorphism originating from variations in the microsatellite repeat length and allows at the same time amplification from multiple microsatellite loci.

All these three methods were shown to be able to distinguish between barley varieties and to produce fingerprint patterns for species across the genus. S-SAP exhibited between one and 39 bands in three barley cultivars using 16 retro-transposon/primer combinations (Waugh et al. 1997). IRAP revealed 100% of polymorphism and displayed approximately 15–30 bands per cultivar. The

amount of polymorphism revealed at the species level within *Hordeum* was even too high to permit the use of IRAP in interspecies phylogenetic analyses (Kalendar et al. 1999). REMAP showed 100% of polymorphism in barley as well. Between 13 and 23 bands were revealed giving more robust patterns compared to IRAP. However, all three techniques produce complex profiles which makes them unsuitable for high throughput genotyping. Moreover, a S-SAP, IRAP or REMAP marker potentially associated to a gene of agronomic interest cannot be efficiently used for physical mapping since the multiple banding pattern will lead to the identification of numerous BAC contigs that will be difficult to individualize.

Three additional methods based on the polymorphic integration pattern of retrotransposons were further developed. The RBIP (Retrotransposon-Based Insertional Polymorphism; Flavell et al. 1998) technique relies on the insertions of a retrotransposon at individual low copy sites (Fig. 9.1D). Primers are designed on both sides of a TE insertion site and used for amplification on different samples. An amplification product indicates the absence of a TE insertion while a lack of amplification suggests its presence. In the latter case, insertion can be confirmed by using a primer located within the transposable element and one of the other two primers. The main limitation of RBIP markers is the large investment needed for their development (Flavell et al. 1998). First, the retrotransposon needs to be characterized prior to any marker development. Databases for the Triticeae are now well documented (TREP: <http://wheat.pw.usda.gov/ITMI/Repeats/>) and the sequences for numerous elements are available making this task easier. Second, the chromosomal region surrounding each TE needs to be cloned and the host insertion flanking site sequenced. This part is more problematic and costly. A microarray-based method has been developed for high throughput scoring of the RBIP markers (Flavell et al. 2003). This approach called tagged microarray marker (TAM) combines streptavidin and hybridisation approaches. Biotinylated allele-specific primers carrying different tags are arrayed onto streptavidin-coated microarray slides which are then hybridised with detector probes recognising tags specific to the two alleles. The TAM approach has been successfully used in barley to score single nucleotide polymorphism (SNP) and to distinguish barley varieties.

A RBIP-derived method, the Insertion Site Based Polymorphism (ISBP) technique, has been described recently (Paux et al. 2006). It was shown to be an efficient way of developing chromosome-specific markers from BAC-end sequences (BES) for genetic and physical mapping in bread wheat. In this technique, two primers flanking the junction between a transposable element and its flanking sequence (either repeated or low copy) are designed and used to amplify specifically the junction sequence (Fig. 9.1E). Out of 20,000 BES, about 3,000 junctions representing various kinds of junctions (two repeated sequences; one repeated sequence/one low-copy sequence; one repeated sequence/one coding sequence) were identified and primers were designed. A chromosome-specific amplification product was obtained in more than

67% of the cases. In addition, more than 53% of the specific markers were polymorphic for at least one of the five wheat lines used in the study, polymorphism being mostly due to the absence of amplification in at least one line. However, size polymorphisms were observed as well and when sequenced, monomorphic bands exhibit numerous SNPs that can be identified with various techniques (Paux et al. unpublished). The main limitation of the technique is the need to produce sequences in which TE insertion sites can be detected. However, since ISBPs can easily be derived from short genomic sequences (e.g. BAC end sequences) that are exponentially growing in the public databases, they represent a very good source of new markers for genetic and physical mapping in the Triticeae.

9.2.5 Diversity Array Technology (DArT) Markers

Diversity arrays technology (DArT) is a privately owned technology (<http://www.diversityarrays.com/>) that was originally developed for rice (*Oryza sativa*), a diploid crop with a small genome of 430 Mbp (Jaccoud et al. 2001). This hybridisation-based method combines the advantages of SNP detection by restriction enzymes and the large capacities of microarrays. DArT allows the generation of whole-genome fingerprints of several thousands of loci in a single assay, by simultaneously scoring the presence versus absence of DNA fragments in genomic representations generated from samples of genomic DNA.

To develop DArTs, diversity panels are generated using DNA from several genotypes representing a good range of variability (Jaccoud et al. 2001) within the target species. Genomic DNA is then digested, ligated with enzyme-specific adapters, amplified and cloned. The cloned fragments are spotted on glass slides to produce the DArT array and hybridized with labelled samples previously prepared following the same selection procedure. The DArT technology was subsequently applied to a range of species including 19 plant species and three fungal plant pathogens (<http://www.diversityarrays.com/>).

DArT arrays have been developed for barley (Wenzl et al. 2004) and wheat (Akbari et al. 2006). In both species, DArT technology performed well and generated a large number of high-quality markers. The genetic relationships among bread wheat or barley cultivars revealed by DArT coincided with knowledge generated with other methods such as RFLPs or SSRs, and allowed distinguishing even between closely related cultivars. Several genetic maps with additional DArT markers were constructed and the coverage was equivalent to what was observed with other types of markers. However, a certain level of clustering was observed (27.0%) at < 0.5 cM intervals between consecutive markers, which is 3 and 15 times higher than SSRs (8.9%) and AFLPs (1.8%), respectively (Hearnden et al. 2007; Semagn et al. 2006). A single DArT assay takes a maximum of three working days to complete from DNA to data and generates a reproducible medium-density scan of the Triticeae genome that is useful for a range of genomics applications, particularly in the context of improving traits with complex inheritance. Recently, DArTs have been

successfully used for QTL mapping in barley (Grewal et al. 2008). In addition, preliminary results suggest that these markers could be useful for physical map anchoring (Kilian A., personal communication).

9.2.6 *Single Nucleotide Polymorphism (SNP) Arrays*

The recent accumulation of data from complete genome sequences, genomic libraries as well as from coding regions with EST sequencing projects has allowed the discovery and the development of Single Nucleotide Polymorphisms (SNPs) based markers in many organisms. SNPs generally behave as codominant markers and can thus be used for projects such as genetic mapping, phylogenetic and structural analyses. Coordinated efforts (ITEC project; <http://wheat.pw.usda.gov/genome/>) resulted in the production of about 1.6 million ESTs from the Triticeae species (<http://www.ncbi.nlm.nih.gov/dbEST/index.html>) and a number of projects have been developed in wheat and barley to use these ESTs for SNP discovery. More than 2000 SNP loci are currently available in wheat (<http://wheat.pw.usda.gov/SNP/new/index.shtml>) and barley (http://bioinf.scri.ac.uk/barley_snpsdb/). SNP discovery is simpler in diploid genomes such as barley (Kanazin et al. 2002; Kota et al. 2001, 2008) than in polyploid species where genome-specific SNPs need to be distinguished from homoeologous copy variations (Somers et al. 2003).

Estimates of SNP frequency (θ) can be computed as the ratio between the number of SNPs and the product between the number of genotypes and the length of the sequence examined (in base pairs; Tenaillon et al. 2001). In maize, the frequency of SNPs is 12.2×10^{-3} ($2.8\text{--}35.6 \times 10^{-3}$; Tenaillon et al. 2001), whereas in wheat, it is half as much (6.9×10^{-3} ; Somers et al. 2003) and even ten times lower in barley (1.1×10^{-3} ; Kanazin et al. 2002). However, these results are based on a few sequence analyses and additional studies are needed to fully understand the distribution of SNPs in Triticeae.

Most SNPs show only two alleles and can be detected with several methods including allele-specific PCR, SNaPshot, SNPlex... More recently, new high-throughput SNP genotyping techniques arose such as the Illumina platform that allows the analysis of 96 to more than 2 million loci in a single assay. In wheat, a 1536 SNP-array has been constructed recently (<http://wheat.pw.usda.gov/SNP/new/index.shtml>). In barley, two projects were established (AGOEUB: <http://www.agoeub.org/> and BarleyCAP: <http://barleycap.coafes.umn.edu/>) resulting in the development of two Illumina Oligo Pool Assay (OPA) based SNP genotyping platforms of 1536 SNPs each that were used for genetic diversity and linkage disequilibrium (LD) analyses, as well as genetic and physical mapping.

Thus, at least six different types of markers comprising several thousands of markers each are now available for wheat and barley and to a lesser extend for rye (Table 9.3). They all have different advantages and drawbacks and can be

Table 9.3 Summary of the characteristics and numbers of the different types of molecular markers available in the Triticeae

Type of Marker	Characteristics		Number of markers			Wheat and related species
	Advantages	Drawbacks	Barley	Rye	Use those of related species.	
RFLP	– Useful on different species – Useful in physical mapping	– Low throughput – Low polymorphism	1,956	No specific markers. Use those of related species.		1,634
SSR	– High throughput – High polymorphism – Chromosome specific	– Expensive to develop	972	27		2,946
AFLP	– High number of combinations – No development needed	– Dominant marker – Difficult to develop STS markers	Unlimited	Unlimited		Unlimited
DArT	– High throughput – High number of markers	– Dominant markers – Private technology	>1,000	Not available yet but can be easily developed		>2,000
ISBP	– Unlimited number – High throughput – High polymorphism	– Dominant markers – Need sequence data	Not developed yet	~100 on 1RS		~1,000 on 3B
SNP	– Unlimited number – High throughput	– Need sequence data – Low number in coding sequences	>2,000	Not developed yet		~4,800

used in combination for genetic and physical mapping, map-based cloning, marker-assisted selection and diversity studies in the Triticeae.

9.3 Expressed Sequence Tag (EST) Sequences and Microarrays

In the absence of a genome sequence, EST sequencing represents an easy way to access the expressed part of a genome. Large efforts have been deployed in the last five years to produce a large set of EST sequences in the Triticeae (wheat, barley, rye and relatives). It resulted in one of the most abundant plant EST collection with almost 1.6 million Triticeae ESTs present in dbEST in August 2008 (<http://www.ncbi.nlm.nih.gov/dbEST/index.html>: with 1,085,694 for wheat and relatives, 502,927 for barley and 9298 for rye). Such a large amount of information has greatly supported Triticeae genomics projects, including comparative genomics (see Chapter 17 for details) and map-based cloning (see Chapter 12 for details). In addition, ESTs have been used for SSR marker development in wheat and barley (see Section 9.2.2). They have also been mapped genetically in barley (Stein et al. 2007) and in deletion bins in wheat (Peng et al. 2004; Conley et al. 2004; Munkvold et al. 2004; Miftahudin et al. 2004; Linkiewicz et al. 2004; Randhawa et al. 2004; Hossain et al. 2004; Qi et al. 2004).

Another important application of ESTs in the last decade has been the development of DNA microarrays. In cereals, microarrays were initially developed in rice (Yazaki et al. 2000) and since then, have been improved and widely used to study the expression of a range of traits such as salt stress (Kawasaki et al. 2001), disease resistance (Shim et al. 2004) or to evaluate the activity of genes related to transposable elements (Jiao and Deng 2007). Moreover, rice microarrays were proven to be useful to study gene expression in related Triticeae species like in barley for iron deficiency (Negishi et al. 2002). Recently an Affymetrix DNA array was constructed for wheat (<http://www.affymetrix.com/products/arrays/specific/wheat.affx>). This array contains 61,127 probes representing 55,052 transcripts for all 21 linkage groups in the wheat genome. The design of the array was based on sequences present in the GenBank and dbEST databases. Sequence information includes public content of the unigene set from *Triticum aestivum* as well as ESTs from the wheat species *T. monococcum*, *T. turgidum*, and *Aegilops tauschii*, and GenBank full-length mRNAs from all closely wheat-related species. This microarray has been used in various studies including developmental processes in wheat (Ogihara et al. 2003), transcript profiling of Near Isogenic Lines for the wheat rust resistance gene Lr34/Yr18 (Hulbert et al. 2007), analysis of the meiosis and microsporogenesis (Crismani et al. 2006), as well as mapping of translocation breakpoints (Bhat et al. 2007). Smaller specific home-made microarrays have also been developed e.g. to compare changes in gene expression in a highly cold-tolerant winter wheat cultivar and a less tolerant spring cultivar (Monroy et al. 2007).

In barley, an Affymetrix DNA array containing 21,349 unigenes and covering about two thirds of the available barley gene sequences was also produced (Close 2005; Close et al. 2004). This microarray was used to study early responses to a salinity stress at the seedling stage in barley (Walia et al. 2006) while smaller arrays were developed to study the expression of genes in specific conditions such as drought and salt stress (Ozturk et al. 2002), cold tolerance (Faccioli et al. 2002) or embryo (Watson and Henry 2005) or anther development (Oshino et al. 2007).

Oligonucleotide arrays can also be used to detect polymorphism in a highly parallel manner (Hazen and Kay 2003). In this case, polymorphism results from differential hybridization of the oligonucleotides present on the array with labeled genomic DNA or mRNA. Single feature polymorphisms (SFPs) that originate from single nucleotide polymorphisms or indels in genomic sequences as well as SNPs present in the transcribed sequence as well as alternative splicing and polyadenylation can be detected using such arrays. Recent studies demonstrated the potential of microarrays for genotyping in the Triticeae. Affymetrix chips were used to detect several thousand SFPs between two barley genotypes (Rostoks et al. 2005; Walia et al. 2007) and to map expression level polymorphism (eQTL) in a wheat cross (Jordan et al. 2007). In total 542 eQTLs were identified that each mapped to a single chromosome interval, illustrating that major gene expression QTLs could be found in wheat.

9.4 Bacterial Artificial Chromosome (BAC) Libraries

Bacterial Artificial Chromosome (BAC) libraries are invaluable tools for genome analysis, physical mapping, map-based cloning and sequencing projects (Shizuya et al. 1992). In the last 15 years, BAC libraries have been constructed from various plant species, including the Triticeae (Table 9.4).

In barley, three BAC libraries were constructed to date, from cultivars Morex (Yu et al. 2000), Cebada Capa (Isidore et al. 2005) and Haruna Nijo (Saisho et al. 2007). The average number of clones for these libraries is about 200,000–300,000, with an average size of 126 kbp and genomic coverage of about 6 genome-equivalents (Table 9.4).

In wheat, BAC libraries were first constructed from diploid relatives. For the A-genome, BAC libraries were developed from *Triticum monococcum*, a cultivated diploid wheat (Lijavetzky et al. 1999) as well as from *T. urartu* which is likely the donor of the A-genome of polyploid wheat (Akhunov et al. 2005; Dvorak et al. 1988; Huang et al. 2002). For the D-genome, three BAC libraries were constructed from *Aegilops tauschii*, the D-genome progenitor of bread wheat (Moulet et al. 1999; Xu et al. 2002; Akhunov et al. 2005). For the B-genome, a single BAC library was reported to date from *Aegilops speltoides* a wild species bearing the S-genome which is closely related to the B-genome donor of wheat (Akhunov et al. 2005; Feldman and Levy 2005). For tetraploid

Table 9.4 List and characteristics of the BAC libraries available for the Triticeae

Plant species	Cultivar	Genome/ chromosome	Cloning site	Number of clones	Insert size	Coverage	References
<i>Hordeum vulgare</i>	Morex	HH	<i>Hind</i> III	313,344	106 kb	6.3 X	Yu et al. 2000
<i>Hordeum vulgare</i>	Cebade Capa	HH	<i>Hind</i> III	177,000	140 kb	6 X	Isidore et al. 2005
<i>Hordeum vulgare</i>	Haruna Nije	HH	<i>Hind</i> III	42,000	115 kb	1.0 X	Saisho et al. 2007
<i>Triticum aestivum</i>	Chinese Spring	AABBDD	<i>Hind</i> III	1,200,000	130 kb	9.3 X	Allouis et al. 2003
<i>Triticum aestivum</i>	Chinese Spring	AABBDD	<i>Hind</i> III	395,136	157 kb	3.4 X	Sher et al. 2005
<i>Triticum aestivum</i>	Renan	AABBDD	<i>Hind</i> III, <i>Eco</i> RI & <i>Bam</i> HI	1,000,320	140 kb	6.9 X	Chalhoub et al., unpublished
<i>Triticum aestivum</i>	Glenlea	AABBDD	<i>Hind</i> III & <i>Bam</i> HI	656,640	80 kb	3.1 X	Nimalgodda et al. 2003
<i>Triticum aestivum</i>	Norstar	AABBDD	<i>Hind</i> III	1,200,000	75 kb	5.5 X	Ratnayaka et al. 2005
<i>Triticum aestivum</i>	AVS-YrS	AABBDD	<i>Hind</i> III	422,400	140 kb	3.6 X	Ling and Chen 2005
<i>Triticum aestivum</i>	Chinese Spring	1D-4D-6D	<i>Hind</i> III	208,512	95 kb	9.7 X	Janda et al. 2004
<i>Triticum aestivum</i>	Chinese Spring	3AS	<i>Hind</i> HI	55,296	80 kb	12.0 X	Dolezel, unpublished
<i>Triticum aestivum</i>	Chinese Spring	3AL	<i>Hind</i> III	55,296	106 kb	12.5 X	Dolezel, unpublished
<i>Triticum aestivum</i>	Chinese Spring	3B	<i>Hind</i> III	67,986	103 kb	6.2 X	Safat et al. 2004
<i>Triticum aestivum</i>	Chinese Spring	3DS	<i>Hind</i> III	36,864	110 kb	11.0 X	Dolezel, unpublished
<i>Triticum aestivum</i>	Chinese Spring	3DL	<i>Hind</i> III	55,296	105 kb	11.2 X	Dolezel, unpublished
<i>Triticum aestivum</i>	Chinese Spring	7DS	<i>Hind</i> III	50,304	115 kb	16.7 X	Dolezel, unpublished

Table 9.4 (continued)

Plant species	Cultivar	Genome/ chromosome	Cloning site	Number of clones	Insert size	Coverage	References
<i>Triticum aestivum</i>	Hope	3B	<i>Hind</i> III	55,296	85 kb	4.5 X	Dolezel, unpublished
<i>Triticum aestivum</i>	Pavon	IBS	<i>Hind</i> III	65,280	82 kb	14.5 X	Janda et al. 2006
<i>Triticum durum</i>	Langdon	AABB	<i>Hind</i> III	516,000	130 kb	5.0 X	Cenci et al. 2003
<i>Triticum monoccum</i>	DV92	AA	<i>Hind</i> III	276,480	115 kb	5.6 X	Lijavetzky et al. 1999
<i>Triticum urartu</i>	G1812	AA	<i>Bam</i> HI	163,200	110 kb	4.9 X	Akhunov et al. 2005
<i>Aegilops tauschii</i>	Ausl8913	DD	<i>Hind</i> III	144,000	119 kb	3.7 X	Moullet et al. 1999
<i>Aegilops tauschii</i>	AL8/78	DD	<i>Hind</i> III, <i>Eco</i> RI & <i>Bam</i> HI	172,000	177 kb	6 X	Xu et al. 2002
<i>Aegilops tauschii</i>	AS75	DD	<i>Bam</i> HI	181,248	115	4.1 X	Akhunov et al. 2005
<i>Aegilops speltoides</i>	2-12-4-8-1- 1-1	SS	<i>Bam</i> HI	237,312	115	5.4 X	Akhunov et al. 2005
<i>Secale cereale</i>	Blanco	RR	<i>Hind</i> III	373,672	130 kb	6 X	Shi et al. 2007
<i>Secale cereale</i>	Imperial	IRS	<i>Hind</i> III & <i>Bam</i> HI	103,680	73 kb	14 X	Dolezel, unpublished

wheat, only one BAC library was reported to date, from the cultivar Langdon (Cenci et al. 2003) whereas several BAC libraries have been constructed for hexaploid wheat (Table 9.4): two from the reference cultivar Chinese Spring (Allouis et al. 2003; Shen et al. 2005), one from cv. Renan (Chalhoub et al. unpublished; www.intl-pag.org/10/abstracts/PAGX_P350.html), one from cv. Glenlea (Nilmalgoda et al. 2003), one from the cv Norstar (Ratnayaka et al. 2005) and one from a BC6 *Yr5* near-isogenic line, originally developed from a cross between cultivar AVS and *T. spelta* 'Album' (the *Yr5* donor; Yan et al. 2003; Ling and Chen 2005).

Because of the large size of the Triticeae genomes (17 Gb for hexaploid, 12 Gb for tetraploid, 5 Gb for the diploid wheat and barley), the BAC libraries consist of a very large number of clones. Indeed, to achieve coverage of one diploid genome-equivalent, approximately 50,000 BAC clones are necessary. As an alternative to whole genome studies, chromosome genomics has been developed in the Triticeae species (see Chapter 10). Recent advances in flow sorting techniques (Kubalakova et al. 2002) have allowed the isolation of DNA in sufficient amounts and quality to construct BAC libraries from single wheat chromosomes or chromosome arms. To date, eight BAC libraries (Table 9.4) have been constructed from the 3B, 3AS, 3AL, 3DS, 3DL, 7DS and 1D-4D-6D chromosomes cv. Chinese Spring (Janda et al. 2004; Safar et al. 2004; Simkova et al. unpublished; Dolezel et al. unpublished; www.ueb.cas.cz/Olomouc1/LMCC/Resources/resources.html) as well as from the short arm of chromosome 1B of hexaploid wheat cv. Pavon (Janda et al. 2006) and from the chromosome 3B of hexaploid wheat cv. Hope (www.ueb.cas.cz/Olomouc1/LMCC/Resources/resources.html; Dolezel et al. unpublished).

The same approach was successfully used in rye where a BAC library specific from the short arm of the chromosome 1R was produced from the cultivar Imperial (Dolezel, unpublished). In addition, a whole-genome BAC library was recently constructed from rye cv. Blanco (Shi et al. 2007) (Table 9.4).

These BAC libraries have been widely used in various projects such as physical mapping (see Chapter 11), map-based cloning of genes (see Chapter 12) and comparative genomics (see Chapter 17).

9.5 Outlook

With the recent development of a toolbox for Triticeae genomics comparable to what is available for rice and maize, Triticeae scientists and breeders can now benefit from a wide range of tools to support genome analyses, as reviewed in the next chapters. Structure and function of the Triticeae genomes can now be accessed, leading to a better understanding of their evolution (see Chapters 17 and 23) as well as to the identification and cloning of genes underlying traits of economic importance such as resistance to biotic and abiotic stresses, plant and grain development or quality traits (see Chapters 18, 19, 20, 21, and Part IV).

Finally, these tools will support future sequencing of the Triticeae genomes (see Chapter 24) that will in turn provide an unlimited number of markers for genetic mapping, molecular breeding and improvement of these essential species.

References

- Akbari, M., Wenzl, P., Caig, V., Carling, J., Xia, L., Yang, S.Y., Uszynski, G., Mohler, V., Lehmensiek, A., Kuchel, H., Hayden, M.J., Howes, N., Sharp, P., Vaughan, P., Rathmell, B., Huttner, E. and Kilian, A. (2006) Diversity arrays technology (DArT) for high-throughput profiling of the hexaploid wheat genome. *Theor. Appl. Genet.* 113, 1409–1420.
- Akhunov, E.D., Akhunova, A.R. and Dvorak, J. (2005) BAC libraries of *Triticum urartu*, *Aegilops speltoides* and *Ae. tauschii*, the diploid ancestors of polyploid wheat. *Theor. Appl. Genet.* 111, 1617–1622.
- Allouis, S., Moore, G., Bellec, A., Sharp, R., Faivre Rampant, P., Mortimer, K., Pateyron, S., Foote, T.N., Griffiths, S., Caboche, M. and Chalhoub, B. (2003) Construction and characterisation of a hexaploid wheat (*Triticum aestivum* L.) BAC library from the reference germplasm ‘Chinese Spring’. *Cereal Res. Commun.* 31, 331–338.
- Appels, R., Francki, M. and Chibbar, R.N. (2003) Advances in cereal functional genomics. *Funct. Integr. Genomics* 3, 1–24.
- Barrett, B.A. and Kidwell, K.K. (1998) AFLP-based genetic diversity assessment among wheat cultivars from the Pacific Northwest. *Crop Sci.* 38, 1261–1271.
- Barrett, B.A., Kidwell, K.K. and Fox, P.N. (1998) Comparison of AFLP and pedigree-based genetic diversity assessment methods using wheat cultivars from the Pacific Northwest. *Crop Sci.* 38, 1271–1278.
- Beaubien, K.A. and Smith, K.P. (2006) New SSR markers for barley derived from the EST database. *Barley Genet. Newsl.* 36, 30–43.
- Becker, J. and Heun, M. (1995) Barley microsatellites-allele variation and mapping. *Plant Mol. Biol.* 27, 835–845.
- Becker, J., Vos, P., Kuiper, M., Salamini, F. and Heun, M. (1995) Combined mapping of AFLP and RFLP markers in barley. *Mol. Gen. Genet.* 249, 65–73.
- Bertin, P., Gregoire, D., Massart, S. and de Froidmont, D. (2001) Genetic diversity among European cultivated spelt revealed by microsatellites. *Theor. Appl. Genet.* 102, 148–156.
- Bhat, P.R., Lukaszewski, A., Cui, X., Xu, J., Svensson, J.T., Wanamaker, S., Waines, J.G. and Close, T.J. (2007) Mapping translocation breakpoints using a wheat microarray. *Nucleic Acids Res.* 35, 2936–2943.
- Blackburn, E.H. (1984) The molecular structure of centromeres and telomeres. *Annu. Rev. Biochem.* 53, 163–194.
- Botstein, D., White, R.L., Skolnick, M. and Davis, R.W. (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.* 32, 314–331.
- Bryan, G., Collins, A., Stephenson, P., Orry, A., Smith, J. and Gale, M.D. (1997) Isolation and characterization of microsatellite from hexaploid bread wheat. *Theor. Appl. Genet.* 94, 557–563.
- Burkhamer, R.L., Lanning, S.P., Martens, R.J., Martin, J.M. and Talbert, L.E. (1998) Predicting progeny variance from parental divergence in hard red spring wheat. *Crop Sci.* 38, 243–248.
- Burr, B., Evola, S., Burr, F. and Beckmann, J. (1983) The application of restriction fragment length polymorphism to plant breeding. In: J. Setlow and A. Hollaender (eds.), *Genetic Engineering: principles and methods*, Plenum Press, New York, pp. 45–59.

- Cadalen, T., Boeuf, C., Bernard, S. and Bernard, M. (1997) An intervarietal molecular marker map in *Triticum aestivum* L. em Thell and comparison with a map from a wide cross. *Theor. Appl. Genet.* 94, 367–377.
- Cenci, A., Chantret, N., Kong, X., Gu, Y., Anderson, O.D., Fahima, T., Distelfeld, A. and Dubcovsky, J. (2003) Construction and characterization of a half million clone BAC library of durum wheat (*Triticum turgidum* ssp. *durum*). *Theor. Appl. Genet.* 107, 931–939.
- Chao, S., Sharp, P.J., Worland, A.J., Warham, E.J., Koeber, R.M.D. and Gale, M.D. (1989) RFLP-based genetic maps of wheat homoeologous group 7 chromosomes. *Theor. Appl. Genet.* 78, 495–504.
- Cheung, W.Y., Moore, G., Money, T.A. and Gale, M.D. (1992) *Hpa*II library indicates methylation-free islands in wheat and barley. *Theor. Appl. Genet.* 84, 739–746.
- Close, T.J. (2005) The barley microarray: a community vision and application to abiotic stress. *Czech J. Genet. Plant Breed.* 41, 144–152.
- Close, T.J., Wanamaker, S.I., Caldo, R.A., Turner, S.M., Ashlock, D.A., Dickerson, J.A., Wing, R.A., Muehlbauer, G.J., Kleinhofs, A. and Wise, R.P. (2004) A new resource for cereal genomics: 22 K barley GeneChip comes of age. *Plant Physiol.* 134, 960–968.
- Conley, E.J., Nduati, V., Gonzalez-Hernandez, J.L., Mesfin, A., Trudeau-Spanjers, M., Chao, S., Lazo, G.R., Hummel, D.D., Anderson, O.D., Qi, L.L., Gill, B.S., Echalié, B., Linkiewicz, A.M., Dubcovsky, J., Akhunov, E.D., Dvorák, J., Peng, J.H., Lapitan, N.L. V., Pathan, M.S., Nguyen, H.T., Ma, X.-F., Miftahudin, Gustafson, J.P., Greene, R.A., Sorrells, M.E., Hossain, K.G., Kalavacharla, V., Kianian, S.F., Sidhu, D., Dilbirli, M., Gill, K.S., Choi, D.W., Fenton, R.D., Close, T.J., McGuire, P.E., Qualset, C.O. and Anderson, J.A. (2004) A 2600-Locus Chromosome Bin Map of Wheat Homoeologous Group 2 Reveals Interstitial Gene-Rich Islands and Colinearity With Rice. *Genetics* 168, 625–637.
- Crismani, W., Baumann, U., Sutton, T., Shirley, N., Webster, T., Spangenberg, G., Langridge, P. and Able, J.A. (2006) Microarray expression analysis of meiosis and microsporogenesis in hexaploid bread wheat. *BMC Genomics* 7, 267.
- Devey, M.E. and Hart, G.E. (1993) Chromosomal localization of intergenomic RFLP loci in hexaploid wheat. *Genome* 36, 913–918.
- Devos, K. and Gale, M. (1993) The genetic maps of wheat and their potential in plant-breeding. *Outlook Agric.* 22, 93–99.
- Devos, K.M., Millan, T. and Gale, M.D. (1993) Comparative RFLP maps of the homoeologous group-2 chromosomes of wheat, rye and barley. *Theor. Appl. Genet.* 85, 784–792.
- Dover, G.A. and Tautz, D. (1986) Conservation and divergence in multigene families: alternatives to selection and drift. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 312, 275–289.
- Dvorak, J., McGuire, P.E. and Cassidy, B. (1988) Apparent sources of the A genomes of wheats inferred from polymorphism in abundance and restriction fragment length of repeated nucleotide sequences. *Genome* 30, 680–689.
- Edwards, K.J., Barker, J.H., Daly, A., Jones, C. and Karp, A. (1996) Microsatellite libraries enriched for several microsatellite sequences in plants. *BioTechniques* 20, 758–760.
- Ellis, R.P., McNicol, J.W., Baird, E., Booth, A., Lawrence, P., Thomas, B. and Powell, W. (1997) The use of AFLPs to examine genetic relatedness in barley. *Mol. Breed.* 3, 359–369.
- Eujayl, I., Sorrells, M.E., Baum, M., Wolters, P. and Powell, W. (2002) Isolation of EST-derived microsatellite markers for genotyping the A and B genomes of wheat. *Theor. Appl. Genet.* 104, 399–407.
- Faccioli, P., Lagonigro, M.S., De Cecco, L., Stanca, A.M., Alberici, R. and Terzi, V. (2002) Analysis of differential expression of barley ESTs during cold acclimatization using microarray technology. *Plant Biol.* 4, 630–639.
- Feldman, M. and Levy, A.A. (2005) Allopolyploidy – a shaping force in the evolution of wheat genomes. *Cytogenet. Genome Res.* 109, 250–258.
- Flavell, A.J., Knox, M.R., Pearce, S.R. and Ellis, T.H. (1998) Retrotransposon-based insertion polymorphisms (RBIP) for high throughput marker analysis. *Plant J.* 16, 643–650.

- Flavell, A.J., Bolshakov, V.N., Booth, A., Jing, R., Russell, J., Ellis, T.H. and Isaac, P. (2003) A microarray-based high throughput molecular marker genotyping method: the tagged microarray marker (TAM) approach. *Nucleic Acids Res.* 31, e115.
- Flavell, R.B., Rimpau, J. and Smith, D.B. (1977) Repeated sequence DNA relationship in four cereal genomes. *Chromosoma* 63, 205–222.
- Gale, M.D., Atkinson, M.D., Chinoy, C.N., Hartcourt, R.L., Jiu, J., Li, Q.Y. and Devos, K.M. (1995) Genetic maps of hexaploid wheat. In: Z.S. Li and Z.Y. Xin (eds.), *Proc 8th Int. Wheat Genet Symp.* Beijing, China, pp. 29–40.
- Gao, L.F., Tang, J.F., Li, H.W. and Jia, J.Z. (2003) Analysis of microsatellites in major crops assessed by computational and experimental approaches. *Mol. Breed.* 12, 245–261.
- Gill, K.S., Lubbers, E.L., Gill, B.S., Raupp, W.J. and Cox, T.S. (1991) A genetic-linkage map of *Triticum tauschii* (DD) and its relationship to the D-genome of bread wheat (AABBDD). *Genome* 34, 362–374.
- Goodman, M. (1985) Rates of molecular evolution: the hominoid slowdown. *BioEssays* 3, 9–14.
- Graner, A., Siedler, H., Jahoor, A., Herrmann, R.G. and Wenzel, G. (1990) Assessment of the degree and the type of Restriction-Fragment-Length-Polymorphism in barley (*Hordeum vulgare*). *Theor. Appl. Genet.* 80, 826–832.
- Graner, A., Jahoor, A., Schondelmaier, J., Siedler, H., Pillen, K., Fischbeck, G., Wenzel, G. and Herrmann, R.G. (1991) Construction of an RFLP map of barley. *Theor. Appl. Genet.* 83, 250–256.
- Greaves, D. and Patient, R. (1985) (AT)_n is an interspersed repeat in the *Xenopus* genome. *EMBO J.* 4, 2617–2626.
- Grewal, T.S., Rosnagel, B.G., Pozniak, C.J. and Scoles, G.J. (2008) Mapping quantitative trait loci associated with barley net blotch resistance. *Theor. Appl. Genet.* 116, 529–539.
- Grodzicker, T., Williams, J., Sharp, P. and Sambrook, J. (1974) Physical mapping of temperature-sensitive mutations of adenoviruses. *Cold Spring Harb. Symp. Quant. Biol.* 39, 439–446.
- Gupta, P.K. and Varshney, R.K. (2000) The development and use of microsatellite markers for genetic analysis and plant breeding with emphasis on bread wheat. *Euphytica* 113, 163–185.
- Gupta, P.K. and Varshney, R. (2004) *Cereal Genomics*. Kluwer Academic Publishers, Dordrecht.
- Gupta, P.K., Rustgi, S., Sharma, S., Singh, R., Kumar, N. and Balyan, H.S. (2003) Transferable EST-SSR markers for the study of polymorphism and genetic diversity in bread wheat. *Mol. Genet. Genomics* 270, 315–323.
- Guyomarc'h, H., Sourdille, P., Edwards, K., Bernard, M. (2002a) Studies of the transferability of microsatellites derived from *Triticum tauschii* to hexaploid wheat and to diploid related species using amplification, hybridization and sequence comparisons. *Theor. Appl. Genet.* 105, 736–744.
- Guyomarc'h, H., Sourdille, P., Charmet, G., Edwards, J. and Bernard, M. (2002b) Characterisation of polymorphic microsatellite markers from *Aegilops tauschii* and transferability to the D-genome of bread wheat. *Theor. Appl. Genet.* 104, 1164–1172.
- Hackauf, B. and Wehling, P. (2002) Identification of microsatellite polymorphisms in an expressed portion of the rye genome. *Plant Breed.* 121, 17–25.
- Hamada, H., Petrino, M.G. and Kakunaga, T. (1982) A novel repeated element with Z-DNA-forming potential is widely found in evolutionarily diverse eukaryotic genomes. *Proc. Natl. Acad. Sci. USA* 79, 6465–6469.
- Hazen, S.P. and Kay, S.A. (2003) Gene arrays are not just for measuring gene expression. *Trends Plant Sci.* 8, 413–416.
- Hearnden, P.R., Eckermann, P.J., McMichael, G.L., Hayden, M.J., Eglinton, J.K. and Chalmers, K.J. (2007) A genetic map of 1,000 SSR and DArT markers in a wide barley cross. *Theor. Appl. Genet.* 115, 383–391.

- Henry, R.J. (2007) Genomics as a tool for cereal chemistry. *Cereal Chem.* 84, 365–369.
- Hentschel, C.C. (1982) Homocopolymer sequences in the spacer of a sea urchin histone gene repeat are sensitive to S1 nuclease. *Nature* 295, 714–716.
- Heun, M., Borghi, B. and Salamini, F. (1998) Wheat domestication – Response. *Science* 279, 303–304.
- Heun, M., Kennedy, A.E., Anderson, J.A., Lapitan, N.L.V., Sorrells, M.E. and Tanksley, S.D. (1991) Construction of a restriction fragment length polymorphism map for barley (*Hordeum vulgare*). *Genome* 34, 437–447.
- Holton, T.A., Christopher, J.T., McClure, L., Harker, N. and Henry, R.J. (2002) Identification and mapping of polymorphic SSR markers from expressed gene sequences of barley and wheat. *Mol. Breed.* 9, 63–71.
- Hossain, K.G., Kalavacharla, V., Lazo, G.R., Hegstad, J., Wentz, M.J., Kianian, P.M.A., Simons, K., Gehlhar, S., Rust, J.L., Syamala, R.R., Obeori, K., Bhamidimarri, S., Karunadharma, P., Chao, S., Anderson, O.D., Qi, L.L., Echali, B., Gill, B.S., Linkiewicz, A.M., Ratnasiri, A., Dubcovsky, J., Akhunov, E.D., Dvorák, J., Miftahudin, Ross, K., Gustafson, J.P., Radhawa, H.S., Dilbirligi, M., Gill, K.S., Peng, J.H., Lapitan, N.L.V., Greene, R.A., Bermudez-Kandianis, C.E., Sorrells, M.E., Feril, O., Pathan, M.S., Nguyen, H.T., Gonzalez-Hernandez, J.L., Conley, E.J., Anderson, J.A., Choi, D.W., Fenton, D., Close, T.J., McGuire, P.E., Qualset, C.O. and Kianian, S.F. (2004) A Chromosome Bin Map of 2148 Expressed Sequence Tag Loci of Wheat Homoeologous Group 7. *Genetics* 168, 687–699.
- Huang, S., Sirikhachornkit, A., Su, X., Faris, J., Gill, B., Haselkorn, R. and Gornicki, P. (2002) Genes encoding plastid acetyl-CoA carboxylase and 3-phosphoglycerate kinase of the *Triticum/Aegilops* complex and the evolutionary history of polyploid wheat. *Proc. Natl. Acad. Sci. USA* 99, 8133–8138.
- Hulbert, S.H., Bai, J., Fellers, J.P., Pacheco, M.G. and Bowden, R.L. (2007) Gene expression patterns in near isogenic lines for wheat rust resistance gene *Lr34/Yr18*. *Phytopathology* 97, 1083–1093.
- Isidore, E., Scherrer, B., Bellec, A., Budin, K., Faivre-Rampant, P., Waugh, R., Keller, B., Caboche, M., Feuillet, C. and Chalhoub, B. (2005) Direct targeting and rapid isolation of BAC clones spanning a defined chromosome region. *Funct. Integr. Genomics* 5, 97–103.
- Jaccoud, D., Peng, K., Feinstein, D. and Kilian, A. (2001) Diversity arrays: a solid state technology for sequence information independent genotyping. *Nucleic Acids Res.* 29, E25.
- Jahoor, A., Backes, G., Graner, A., Herrmann, R.G. and Fischbeck, G. (1991) Development of RFLP markers for barley. *Plant Breed.* 107, 73–76.
- Janda, J., Bartos, J., Safar, J., Kubalaková, M., Valarik, M., Cihalikova, J., Simkova, H., Caboche, M., Sourdille, P., Bernard, M., Chalhoub, B. and Dolezel, J. (2004) Construction of a subgenomic BAC library specific for chromosomes 1D, 4D and 6D of hexaploid wheat. *Theor. Appl. Genet.* 109, 1337–1345.
- Janda, J., Safar, J., Kubalaková, M., Bartos, J., Kovarova, P., Suchankova, P., Pateyron, S., Cihalikova, J., Sourdille, P., Simkova, H., Faivre-Rampant, P., Hribova, E., Bernard, M., Lukaszewski, A., Dolezel, J. and Chalhoub, B. (2006) Advanced resources for plant genomics: a BAC library specific for the short arm of wheat chromosome 1B. *Plant J.* 47, 977–986.
- Jeffreys, A.J., Wilson, V. and Thein, S.L. (1985) Hypervariable ‘minisatellite’ regions in human DNA. *Nature* 314, 67–73.
- Jiao, Y. and Deng, X.W. (2007) A genome-wide transcriptional activity survey of rice transposable element-related genes. *Genome Biol.* 8, R28.
- Jordan, M.C., Somers, D.J. and Banks, T.W. (2007) Identifying regions of the wheat genome controlling seed development by mapping expression quantitative trait loci. *Plant Biotechnol. J.* 5, 442–453.
- Kalendar, R., Grob, T., Regina, M., Suoniemi, A. and Schulman, A. (1999) IRAP and REMAP: two new retrotransposon-based DNA fingerprinting techniques. *Theor. Appl. Genet.* 98, 704–711.

- Kammorgan, L.N.W., Gill, B.S. and Muthukrishnan, S. (1989) DNA restriction fragment length polymorphisms – a strategy for genetic-mapping of D genome of wheat. *Genome* 32, 724–732.
- Kanazin, V., Talbert, H., See, D., DeCamp, P., Nevo, E. and Blake, T. (2002) Discovery and assay of single-nucleotide polymorphisms in barley (*Hordeum vulgare*). *Plant Mol. Biol.* 48, 529–537.
- Kantety, R.V., La Rota, M., Matthews, D.E. and Sorrells, M.E. (2002) Data mining for simple sequence repeats in expressed sequence tags from barley, maize, rice, sorghum and wheat. *Plant Mol. Biol.* 48, 501–510.
- Kawasaki, S., Borchert, C., Deyholos, M., Wang, H., Brazille, S., Kawai, K., Galbraith, D. and Bohnert, H.J. (2001) Gene expression profiles during the initial phase of salt stress in rice. *Plant Cell* 13, 889–905.
- Keim, P., Diers, B.W., Olson, T.C. and Shoemaker, R.C. (1990) RFLP mapping in soybean: association between marker loci and variation in quantitative traits. *Genetics* 126, 735–742.
- Kleinhofs, A., Kilian, A., Maroof, M.A.S., Biyashev, R.M., Hayes, P., Chen, F.Q., Lapitan, N., Fenwick, A., Blake, T.K., Kanazin, V., Ananiev, E., Dahleen, L., Kudrna, D., Bollinger, J., Knapp, S.J., Liu, B., Sorrells, M., Heun, M., Franckowiak, J.D., Hoffman, D., Skadsen, R. and Steffenson, B.J. (1993) A molecular, isozyme and morphological map of the barley (*Hordeum vulgare*) genome. *Theor. Appl. Genet.* 86, 705–712.
- Kota, R., Varshney, R.K., Thiel, T., Dehmer, K.J. and Graner, A. (2001) Generation and comparison of EST-derived SSRs and SNPs in barley (*Hordeum vulgare* L.). *Hereditas* 135, 145–151.
- Kota, R., Varshney, R.K., Prasad, M., Zhang, H., Stein, N. and Graner, A. (2008) EST-derived single nucleotide polymorphism markers for assembling genetic and physical maps of the barley genome. *Funct. Integr. Genomics* 8, 223–233.
- Kubalakova, M., Vrana, J., Cihalikova, J., Simkova, H. and Dolezel, J. (2002) Flow karyotyping and chromosome sorting in bread wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* 104, 1362–1372.
- Leigh, F., Lea, V., Law, J., Wolters, P., Powell, W. and Donini, P. (2003) Assessment of EST- and genomic microsatellite markers for variety discrimination and genetic diversity studies in wheat. *Euphytica* 133, 359–366.
- Li, W., Zhang, P., Fellers, J.P., Friebe, B. and Gill, B.S. (2004) Sequence composition, organization, and evolution of the core Triticeae genome. *Plant J.* 40, 500–511.
- Lijavetzky, D., Muzzi, G., Wicker, T., Keller, B., Wing, R. and Dubcovsky, J. (1999) Construction and characterization of a bacterial artificial chromosome (BAC) library for the A genome of wheat. *Genome* 42, 1176–1182.
- Ling, P. and Chen, X.M. (2005) Construction of a hexaploid wheat (*Triticum aestivum* L.) bacterial artificial chromosome library for cloning genes for stripe rust resistance. *Genome* 48, 1028–1036.
- Linkiewicz, A.M., Qi, L.L., Gill, B.S., Ratnasiri, A., Echalié, B., Chao, S., Lazo, G.R., Hummel, D.D., Anderson, O.D., Akhunov, E.D., Dvorač, J., Pathan, M.S., Nguyen, H. T., Peng, J.H., Lapitan, N.L.V., Miftahudin, Gustafson, J.P., La Rota, C.M., Sorrells, M. E., Hossain, K.G., Kalavacharla, V., Kianian, S.F., Sandhu, D., Bondareva, S.N., Gill, K. S., Conley, E.J., Anderson, J.A., Fenton, R.D., Close, T.J., McGuire, P.E., Qualset, C.O. and Dubcovsky, J. (2004) A 2500-Locus Bin Map of Wheat Homoeologous Group 5 Provides Insights on Gene Distribution and Colinearity With Rice. *Genetics* 168, 665–676.
- Liu, Y.G. and Tsunewaki, K. (1991) Restriction fragment length polymorphism (RFLP) analysis in wheat. II. Linkage maps of the RFLP sites in common wheat. *Jpn. J. Genet.* 66, 617–633.
- Marino, C., Nelson, J., Lu, Y., Sorrells, M., Leroy, P., Lopes, C. and Hart, G. (1996) RFLP-based linkage maps of the homoeologous group 6 chromosomes of hexaploid wheat (*Triticum aestivum* L. em. Thell). *Genome* 39, 359–366.

- Metzgar, D., Bytof, J. and Wills, C. (2000) Selection against frameshift mutations limits microsatellite expansion in coding DNA. *Genome Res.* 10, 72–80.
- Miftahudin, Ross, K., Ma, X.-F., Mahmoud, A.A., Layton, J., Rodriguez Milla, M.A., Chikmawati, T., Ramalingam, J., Feril, O., Pathan, M.S., Surlan Momirovic, G., Kim, S., Chema, K., Fang, P., Haule, L., Struxness, H., Birkes, J., Yaghoubian, C., Skinner, R., McAllister, J., Nguyen, V., Qi, L.L., Echalié, B., Gill, B.S., Linkiewicz, A.M., Dubcovsky, J., Akhunov, E.D., Dvorač, J., Dilbirligi, M., Gill, K.S., Peng, J.H., Lapitan, N.L.V., Bermudez-Kandianis, C.E., Sorrells, M.E., Hossain, K.G., Kalavacharla, V., Kianian, S.F., Lazo, G.R., Chao, S., Anderson, O.D., Gonzalez-Hernandez, J., Conley, E.J., Anderson, J.A., Choi, D.-W., Fenton, R.D., Close, T.J., McGuire, P.E., Qualset, C.O., Nguyen, H.T. and Gustafson, J.P. (2004) Analysis of Expressed Sequence Tag Loci on Wheat Chromosome Group 4. *Genetics* 168, 651–663.
- Mingeot, D. and Jacquemin, J.M. (1999) Mapping of RFLP probes characterized for their polymorphism on wheat. *Theor. Appl. Genet.* 98, 1132–1137.
- Monroy, A.F., Dryanova, A., Malette, B., Oren, D.H., Ridha Farajalla, M., Liu, W., Danyluk, J., Ubayasena, L.W., Kane, K., Scoles, G.J., Sarhan, F. and Gulick, P.J. (2007) Regulatory gene candidates and gene expression analysis of cold acclimation in winter and spring wheat. *Plant Mol. Biol.* 64, 409–423.
- Moore, G., Devos, K.M., Wang, Z. and Gale, M.D. (1995) Cereal genome evolution. Grasses, line up and form a circle. *Curr. Biol.* 5, 737–739.
- Morgante, M., Hanafey, M. and Powell, W. (2002) Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes. *Nat. Genet.* 30, 194–200.
- Moulet, O., Zhang, H.B. and Lagudah, E.S. (1999) Construction and characterisation of a large DNA insert library from the D genome of wheat. *Theor. Appl. Genet.* 99, 305–313.
- Muniz, L.M., Cuadrado, A., Jouve, N. and Gonzalez, J.M. (2001) The detection, cloning, and characterisation of WIS 2–1A retrotransposon-like sequences in *Triticum aestivum* L. and *Triticosecale* Wittmack and an examination of their evolution in related Triticeae. *Genome* 44, 979–989.
- Munkvold, J.D., Greene, R.A., Bermudez-Kandianis, C.E., La Rota, C.M., Edwards, H., Sorrells, S.F., Dake, T., Benscher, D., Kantety, R., Linkiewicz, A.M., Dubcovsky, J., Akhunov, E.D., Dvorač, J., Miftahudin, Gustafson, J.P., Pathan, M.S., Nguyen, H.T., Matthews, D.E., Chao, S., Lazo, G.R., Hummel, D.D., Anderson, O.D., Anderson, J.A., Gonzalez-Hernandez, J.L., Peng, J.H., Lapitan, N., Qi, L.L., Echalié, B., Gill, B.S., Hossain, K.G., Kalavacharla, V., Kianian, S.F., Sandhu, D., Erayman, M., Gill, K.S., McGuire, P.E., Qualset, C.O. and Sorrells, M.E. (2004) Group 3 Chromosome Bin Maps of Wheat and Their Relationship to Rice Chromosome 1. *Genetics* 168, 639–650.
- Negishi, T., Nakanishi, H., Yazaki, J., Kishimoto, N., Fujii, F., Shimbo, K., Yamamoto, K., Sakata, K., Sasaki, T., Kikuchi, S., Mori, S. and Nishizawa, N.K. (2002) cDNA microarray analysis of gene expression during Fe-deficiency stress in barley suggests that polar transport of vesicles is implicated in phytosiderophore secretion in Fe-deficient barley roots. *Plant J.* 30, 83–94.
- Nelson, J.C., Vandeynze, A.E., Autrique, E., Sorrells, M.E., Lu, Y.H., Negre, S., Bernard, M. and Leroy, P. (1995a) Molecular mapping of wheat – homoeologous group-3. *Genome* 38, 525–533.
- Nelson, J.C., Vandeynze, A.E., Autrique, E., Sorrells, M.E., Lu, Y.H., Merlino, M., Atkinson, M. and Leroy, P. (1995b) Molecular mapping of wheat – homoeologous group-2. *Genome* 38, 516–524.
- Nelson, J.C., Sorrells, M.E., Vandeynze, A.E., Lu, Y.H., Atkinson, M., Bernard, M., Leroy, P., Faris, J.D. and Anderson, J.A. (1995c) Molecular mapping of wheat – major genes and rearrangements in homoeologous group-4, group-5, and group-7. *Genetics* 141, 721–731.
- Nicot, N., Chiquet, V., Gandon, B., Amilhat, L., Legeai, F., Leroy, P., Bernard, M. and Sourdille, P. (2004) Study of simple sequence repeat (SSR) markers from wheat expressed sequence tags (ESTs). *Theor. Appl. Genet.* 109, 800–805.

- Nilmalgoda, S.D., Cloutier, S. and Walichnowski, A.Z. (2003) Construction and characterization of a bacterial artificial chromosome (BAC) library of hexaploid wheat (*Triticum aestivum* L.) and validation of genome coverage using locus-specific primers. *Genome* 46, 870–878.
- Ogihara, Y., Mochida, K., Nemoto, Y., Murai, K., Yamazaki, Y., Shin, I.T. and Kohara, Y. (2003) Correlated clustering and virtual display of gene expression patterns in the wheat life cycle by large-scale statistical analyses of expressed sequence tags. *Plant J.* 33, 1001–1011.
- Oshino, T., Abiko, M., Saito, R., Ichiishi, E., Endo, M., Kawagishi-Kobayashi, M. and Higashitani, A. (2007) Premature progression of anther early developmental programs accompanied by comprehensive alterations in transcription during high-temperature injury in barley plants. *Mol. Genet. Genomics* 278, 31–42.
- Ostrander, E.A., Jong, P.M., Rine, J. and Duyk, G. (1992) Construction of small-insert genomic DNA libraries highly enriched for microsatellite repeat sequences. *Proc. Natl Acad. Sci. USA* 89, 3419–3423.
- Ozturk, Z.N., Talame, V., Deyholos, M., Michalowski, C.B., Galbraith, D.W., Gozukirmizi, N., Tuberosa, R. and Bohnert, H.J. (2002) Monitoring large-scale changes in transcript abundance in drought- and salt-stressed barley. *Plant Mol. Biol.* 48, 551–573.
- Pakniyat, H., Powell, W., Baird, E., Handley, L.L., Robinson, D., Scrimgeour, C.M., Nevo, E., Hackett, C.A., Caligari, P.D.S. and Forster, B.P. (1997) AFLP variation in wild barley (*Hordeum spontaneum* C. Koch) with reference to salt tolerance and associated ecogeography. *Genome* 40, 332–341.
- Panstruga, R., Buschges, R., Piffanelli, P. and Schulze-Lefert, P. (1998) A contiguous 60 kb genomic stretch from barley reveals molecular evidence for gene islands in a monocot genome. *Nucleic Acids Res.* 26, 1056–1062.
- Paux, E., Roger, D., Badaeva, E., Gay, G., Bernard, M., Sourdille, P. and Feuillet, C. (2006) Characterizing the composition and evolution of homoeologous genomes in hexaploid wheat through BAC-end sequencing on chromosome 3B. *Plant J.* 48, 463–474.
- Peng, J.H., Zadeh, H., Lazo, J.P., Gustafson, J.P., Chao, S., Anderson, O.D., Qi, L.L., Echalié, B., Gill, B.S., Dilbirligi, M., Gill, K.S., Greene, R.A., Sorrells, M.E., Akhunov, E.D., Dvorač, J., Linkiewicz, A.M., Dubcovsky, J., Hossain, K.G., Kalavacharla, V., Kianian, S.F., Mahmoud, A.A., Miftahudin, Conley, E.J., Anderson, J.A., Pathan, M.S., Nguyen, H.T., McGuire, P.E., Qualset, C.O. and Lapitan, N.L.V. (2004) Chromosome Bin Map of Expressed Sequence Tags in Homoeologous Group 1 of Hexaploid Wheat and Homoeology With Rice and Arabidopsis. *Genetics* 168, 609–623.
- Pestsova, E., Ganal, M.W. and Röder, M.S. (2000) Isolation and mapping of microsatellite markers specific for the D genome of bread wheat. *Genome* 43, 689–697.
- Pillen, K., Binder, A., Kreuzkam, B., Ramsay, L., Waugh, R., Förster, J. and Léon, J. (2000) Mapping new EMBL-derived barley microsatellites and their use in differentiating German barley cultivars. *Theor. Appl. Genet.* 101, 652–660.
- Powell, W., Thomas, W.T.B., Baird, E., Lawrence, P., Booth, A., Harrower, B., McNicol, J. W. and Waugh, R. (1997) Analysis of quantitative traits in barley by the use of amplified fragment length polymorphisms. *Heredity* 79, 48–59.
- Qi, X. and Lindhout, P. (1997) Development of AFLP markers in barley. *Mol. Gen. Genet.* 254, 330–336.
- Qi, L.L., Echalié, B., Chao, S., Lazo, G.R., Butler, G.E., Anderson, O.D., Akhunov, E.D., Dvorač, J., Linkiewicz, A.M., Ratnasiri, A., Dubcovsky, J., Bermudez-Kandianis, C.E., Greene, R.A., Kantety, R., La Rota, C.M., Munkvold, J.D., Sorrells, S.F., Sorrells, M.E., Dilbirligi, M., Sidhu, D., Erayman, M., Randhawa, H.S., Sandhu, D., Bondareva, S.N., Gill, K.S., Mahmoud, A.A., Ma, X.-F., Miftahudin, Gustafson, J.P., Conley, E.J., Nduati, V., Gonzalez-Hernandez, J.L., Anderson, J.A., Peng, J.H., Lapitan, N.L.V., Hossain, K.G., Kalavacharla, V., Kianian, S.F., Pathan, M.S., Zhang, D.S., Nguyen, H.T., Choi, D.-W., Fenton, R.D., Close, T.J., McGuire, P.E., Qualset, C.O. and Gill, B.S.

- (2004) A Chromosome Bin Map of 16,000 Expressed Sequence Tag Loci and Distribution of Genes Among the Three Genomes of Polyploid Wheat. *Genetics* 168, 701–712.
- Quarrie, S.A., Steed, A., Calestani, C., Semikhodskii, A., Lebreton, C., Chinoy, C., Steele, N., Pljevljakusić, D., Waterman, E., Weyen, J., Schondelmaier, J., Habash, D.Z., Farmer, P., Saker, L., Clarkson, D.T., Abugalieva, A., Yessimbekova, M., Turuspekov, Y., Abugalieva, S., Tuberosa, R., Sanguineti, M.C., Hollington, P.A., Aragues, R., Royo, A. and Dodig, D. (2005) A high-density genetic map of hexaploid wheat (*Triticum aestivum* L.) from the cross Chinese Spring × SQ1 and its use to compare QTLs for grain yield across a range of environments. *Theor. Appl. Genet.* 110, 865–880.
- Ramsay, L., Macaulay, M., degli Ivanissevich, S., MacLean, K., Cardle, L., Fuller, J., Edwards, K.J., Tuvešson, S., Morgante, M., Massari, A., Maestri, E., Marmioli, N., Sjakste, T., Ganai, M., Powell, W. and Waugh, R. (2000) A simple sequence repeat-based linkage map of barley. *Genetics* 156, 1997–2005.
- Randhawa, H.S., Dilbirligi, M., Sidhu, D., Erayman, M., Sandhu, D., Bondareva, S., Chao, S., Lazo, G.R., Anderson, O.D., Miftahudin, Gustafson, J.P., Echalié, B., Qi, L.L., Gill, B.S., Akhunov, E.D., Dvorač, J., Linkiewicz, A.M., Ratnasiri, A., Dubcovsky, J., Bermudez-Kandianis, C.E., Greene, R.A., Sorrells, M.E., Conley, E.J., Anderson, J.A., Peng, J.H., Lapitan, N.L.V., Hossain, K.G., Kalavacharla, V., Kianian, S.F., Pathan, M.S., Nguyen, H.T., Endo, T.R., Close, T.J., McGuire, P.E., Qualset, C.O. and Gill, K.S. (2004) Deletion Mapping of Homoeologous Group 6-Specific wheat Expressed Sequence Tags. *Genetics* 168, 677–686.
- Ratnayaka, I., Baga, M., Fowler, D.B. and Chibbar, R.N. (2005) Construction and characterization of a BAC library of a cold-tolerant hexaploid wheat cultivar. *Crop Sci.* 45, 1571–1577.
- Röder, M.S., Plaschke, J., König, S.U., Börner, A., Sorrells, M.E., Tanksley, S.D. and Ganai, M.W. (1995) Abundance, variability and chromosomal location of microsatellites in wheat. *Mol. Gen. Genet.* 246, 327–333.
- Röder, M.S., Korzun, V., Wendehake, K., Plaschke, J., Tixier, M.H., Leroy, P. and Ganai, M.W. (1998) A microsatellite map of wheat. *Genetics* 149, 2007–2023.
- Rostoks, N., Park, Y.J., Ramakrishna, W., Ma, J.X., Druka, A., Shiloff, B.A., SanMiguel, P. J., Jiang, Z., Bruggeman, R., Sandhu, D., Gill, K., Bennetzen, J.L. and Kleinhofs, A. (2002) Genomic sequencing reveals gene content, genomic organization, and recombination relationships in barley. *Func. Int. Genomics* 2, 51–59.
- Rostoks, N., Borevitz, J.O., Hedley, P.E., Russell, J., Mudie, S., Morris, J., Cardle, L., Marshall, D.F. and Waugh, R. (2005) Single-feature polymorphism discovery in the barley transcriptome. *Genome Biol.* 6, R54.
- Russell, J., Fuller, J., Young, G., Thomas, B., Taramino, G., Macaulay, M., Waugh, R. and Powell, W. (1997) Discriminating between barley genotypes using microsatellite markers. *Genome* 40, 442–450.
- Saal, B. and Wricke, G. (1999) Development of simple sequence repeat markers in rye (*Secale cereale* L.). *Genome* 42, 964–972.
- Sabot, F., Guyot, R., Wicker, T., Chantret, N., Laubin, B., Chalhou, B., Leroy, P., Sourdille, P. and Bernard, M. (2005) Updating of transposable element annotations from large wheat genomic sequences reveals diverse activities and gene associations. *Mol. Genet. Genomics* 274, 119–130.
- Safar, J., Bartos, J., Janda, J., Bellec, A., Kubalakova, M., Valarik, M., Pateyron, S., Weiserova, J., Tuskova, R., Cihalikova, J., Vrana, J., Simkova, H., Faivre-Rampant, P., Sourdille, P., Caboche, M., Bernard, M., Dolezel, J. and Chalhou, B. (2004) Dissecting large and complex genomes: flow sorting and BAC cloning of individual chromosomes from bread wheat. *Plant J.* 39, 960–968.
- Saisho, D., Myoraku, E., Kawasaki, S., Sato, K. and Takeda, K. (2007) Construction and Characterization of a Bacterial Artificial Chromosome (BAC) Library from the Japanese Malting Barley variety ‘Haruna Nijo’. *Breed. Sci.* 57, 29–38.

- Schut, J.W., Qi, X. and Stam, P. (1997) Association between relationship measures based on AFLP markers, pedigree data and morphological traits in barley. *Theor. Appl. Genet.* 95, 1161–1168.
- Semagn, K., Bjørnstad, A., Skjinner, H., Maroy, A.G., Tarkegne, Y. and William, M. (2006) Distribution of DArT, AFLP, and SSR markers in a genetic linkage map of a doubled-haploid hexaploid wheat population. *Genome* 49, 545–555.
- Shan, X., Blake, T.K. and Talbert, L.E. (1999) Conversion of AFLP markers to sequence-specific PCR markers in barley and wheat. *Theor. Appl. Genet.* 98, 1072–1078.
- Sharp, P.J., Chao, S., Desai, S. and Gale, M.D. (1989) The isolation, characterization and application in the Triticeae of a set of wheat RFLP probes identifying each homoeologous chromosome arm. *Theor. Appl. Genet.* 78, 342–348.
- Shen, B., Wang, D.M., McIntyre, C.L. and Liu, C.J. (2005) A ‘Chinese Spring’ wheat (*Triticum aestivum* L.) bacterial artificial chromosome library and its use in the isolation of SSR markers for targeted genome regions. *Theor. Appl. Genet.* 111, 1489–1494.
- Shi, B.-J., Miftahudin, Collins, N., Langridge, P. and Gustafson, J.P. (2007) Construction of a rye cv. Blanco BAC library, and progress towards cloning the rye Alt3 aluminium [aluminum] tolerance gene. *Vortr. Pflanzenzüchtg.* 71, 205–209.
- Shim, K.S., Cho, S.K., Jeung, J.U., Jung, K.W., You, M.K., Ok, S.H., Chung, Y.S., Kang, K. H., Hwang, H.G., Choi, H.C., Moon, H.P. and Shin, J.S. (2004) Identification of fungal (*Magnaporthe grisea*) stress-induced genes in wild rice (*Oryza minuta*). *Plant Cell Rep.* 22, 599–607.
- Shizuya, H., Birren, B., Kim, U.J., Mancino, V., Slepak, T., Tachiiri, Y. and Simon, M. (1992) Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor-based vector. *Proc. Natl Acad. Sci. USA* 89, 8794–8797.
- Somers, D.J., Kirkpatrick, R., Moniwa, M. and Walsh, A. (2003) Mining single-nucleotide polymorphisms from hexaploid wheat ESTs. *Genome* 46, 431–437.
- Somers, D.J., Isaac, P. and Edwards, K. (2004) A high-density microsatellite consensus map for bread wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* 109, 1105–1114.
- Stachel, M., Lelley, T., Grausgruber, H. and Vollmann, J. (2000) Application of microsatellites in wheat (*Triticum aestivum* L.) for studying genetic differentiation caused by selection for adaptation and use. *Theor. Appl. Genet.* 100, 242–248.
- Stein, N., Prasad, M., Scholz, U., Thiel, T., Zhang, H.N., Wolf, M., Kota, R., Varshney, R.K., Perovic, D., Grosse, I. and Graner, A. (2007) A 1,000-loci transcript map of the barley genome: new anchoring points for integrative grass genomics. *Theor. Appl. Genet.* 114, 823–839.
- Struss, D. and Pliescke, J. (1998) The use of microsatellite markers for detection of genetic diversity in barley populations. *Theor. Appl. Genet.* 97, 308–315.
- Suoniemi, A., Ananthawat-Jonsson, K., Arna, T. and Schulman, A.H. (1996) Retrotransposon BARE-1 is a major, dispersed component of the barley (*Hordeum vulgare* L.) genome. *Plant Mol. Biol.* 30, 1321–1329.
- Taketa, S., Ando, H., Takeda, K., Harrison, G.E. and Heslop-Harrison, J.S. (2000) The distribution, organization and evolution of two abundant and widespread repetitive DNA sequences in the genus *Hordeum*. *Theor. Appl. Genet.* 100, 169–176.
- Tanksley, S.D., Miller, J., Paterson, A. and Bernatzky, R. (1988) Molecular mapping of plant chromosomes. In: J.F. Gustafson and R. Appels (eds.), *Chromosome structure and function*, Plenum Press, New York, pp. 157–173.
- Tautz, D. (1993) Notes on the definition and nomenclature of tandemly repetitive DNA sequences. *Exs* 67, 21–28.
- Tautz, D. and Renz, M. (1984) Simple sequences are ubiquitous repetitive components of eukaryotic genomes. *Nucleic Acids Res.* 12, 4127–4138.
- Tautz, D., Trick, M. and Dover, G.A. (1986) Cryptic simplicity in DNA is a major source of genetic variation. *Nature* 322, 652–656.

- Tenaillon, M.I., Sawkins, M.C., Long, A.D., Gaut, R.L., Doebley, J.F. and Gaut, B.S. (2001) Patterns of DNA sequence polymorphism along chromosome 1 of maize (*Zea mays* ssp. *mays* L.). *Proc. Natl. Acad. Sci. USA* 98, 9161–9166.
- Thiel, T., Michalek, W., Varshney, R.K. and Graner, A. (2003) Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). *Theor. Appl. Genet.* 106, 411–422.
- Varshney, R.K., Grosse, I., Hahnel, U., Siefken, R., Prasad, M., Stein, N., Langridge, P., Altschmied, L. and Graner, A. (2006a) Genetic mapping and BAC assignment of EST-derived SSR markers shows nonuniform distribution of genes in the barley genome. *Theor. Appl. Genet.* 113, 239–250.
- Varshney, R.K., Hoisington, D.A. and Tyagi, A.K. (2006b) Advances in cereal genomics and applications in crop breeding. *Trends Biotechnol.* 24, 490–499.
- Varshney, R.K., Thiel, T., Stein, N., Langridge, P. and Graner, A. (2002) *In silico* analysis on frequency and distribution of microsatellites in ESTs of some cereal species. *Cell. Mol. Biol. Lett.* 7, 537–546.
- Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M. and Zabeau, M. (1995) AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res.* 23, 4407–4414.
- Walia, H., Wilson, C., Wahid, A., Condamine, P., Cui, X. and Close, T.J. (2006) Expression analysis of barley (*Hordeum vulgare* L.) during salinity stress. *Funct. Integr. Genomics* 6, 143–156.
- Walia, H., Wilson, C., Condamine, P., Ismail, A.M., Xu, J., Cui, X. and Close, T.J. (2007) Array-based genotyping and expression analysis of barley cv. Maythorpe and Golden Promise. *BMC Genomics* 8, 87.
- Wang, A.H., Quigley, G.J., Kolpak, F.J., Crawford, J.L., van Boom, J.H., van der Marel, G. and Rich, A. (1979) Molecular structure of a left-handed double helical DNA fragment at atomic resolution. *Nature* 282, 680–686.
- Watson, L. and Henry, R.J. (2005) Microarray analysis of gene expression in germinating barley embryos (*Hordeum vulgare* L.). *Funct. Integr. Genomics* 5, 155–162.
- Waugh, R., McLean, K., Flavell, A.J., Pearce, S.R., Kumar, A., Thomas, B.B. and Powell, W. (1997) Genetic distribution of BARE-1-like retrotransposable elements in the barley genome revealed by sequence-specific amplification polymorphisms (S-SAP). *Mol. Gen. Genet.* 253, 687–694.
- Weber, J.L. (1990) Informativeness of human (dC-dA)_n(dG-dT)_n polymorphisms. *Genomics* 7, 524–530.
- Wei, F., Wing, R.A. and Wise, R.P. (2002) Genome dynamics and evolution of the *Mla* (powdery mildew) resistance locus in barley. *Plant Cell* 14, 1903–1917.
- Weintraub, H. and Groudine, M. (1976) Chromosomal subunits in active genes have an altered conformation. *Science* 193, 848–856.
- Wenzl, P., Carling, J., Kudrna, D., Jaccoud, D., Huttner, E., Kleinbols, A. and Kilian, A. (2004) Diversity Arrays Technology (DART) for whole-genome profiling of barley. *Proc. Natl. Acad. Sci. USA* 101, 9915–9920.
- Xu, Z., Deal, K., Li, W., Covaleda, L., Chang, Y.-L., Dvorak, J., Luo, M.-C., Gill, B., Anderson, O. and Zhang, H. (2002) Construction and characterization of five large-insert BAC and BIBAC libraries of *Aegilops tauschii*, the diploid donor of the wheat D genome. In: *10th international plant and animal genome conference*, San Diego, USA, pp. 101.
- Yan, G.P., Chen, X.M., Line, R.F. and Wellings, C. R. (2003) Resistance gene-analog polymorphism markers co-segregating with the *Yr5* gene for resistance to wheat stripe rust. *Theor. Appl. Genet.* 106, 636–643.
- Yazaki, J., Kishimoto, N., Nakamura, K., Fujii, F., Shimbo, K., Otsuka, Y., Wu, J., Yamamoto, K., Sakata, K., Sasaki, T. and Kikuchi, S. (2000) Embarking on rice functional genomics via cDNA microarray: use of 3' UTR probes for specific gene expression analysis. *DNA Res.* 7, 367–370.

- Yu, Y., Tomkins, J.P., Waugh, R., Frisch, D.A., Kudrna, D., Kleinhofs, A., Brueggeman, R. S., Muehlbauer, G.J., Wise, R.P. and Wing, R.A. (2000) A bacterial artificial chromosome library for barley (*Hordeum vulgare* L.) and the identification of clones containing putative resistance genes. *Theor. Appl. Genet.* 101, 1093–1099.
- Zabeau, M. and Vos, P. (1993) Selective restriction fragment amplification: a general method for DNA fingerprinting. European Patent Application No. 92402629.7. Publication No. 0534858.
- Zhang, L.Y., Bernard, M., Leroy, P., Feuillet, C. and Sourdille, P. (2005) High transferability of bread wheat EST-derived SSRs to other cereals. *Theor. Appl. Genet.* 111, 677–687.
- Zhang, L.Y., Ravel, C., Bernard, M., Balfourier, F., Leroy, P., Feuillet, C. and Sourdille, P. (2006) Transferable bread wheat EST-SSRs can be useful for phylogenetic studies among the Triticeae species. *Theor. Appl. Genet.* 113, 407–418.
- Zhang, L.Y., Bernard, M., Ravel, C., Balfourier, F., Leroy, P., Feuillet, C. and Sourdille, P. (2007) Wheat EST-SSRs for tracing chromosome segments from a wide range of grass species. *Plant Breed.* 126, 251–258.
- Zonneveld, B.J., Leitch, I.J. and Bennett, M.D. (2005) First nuclear DNA amounts in more than 300 angiosperms. *Ann. Bot. (Lond.)* 96, 229–244.

Chapter 11

Physical Mapping in the Triticeae

Nils Stein

Abstract In contrast to small genome model species where whole genome “shotgun sequencing” is sufficient, physical maps are mandatory for the development of whole genome reference sequences of large and complex genomes, such as those of the Triticeae crop species wheat, barley, and rye. Access to a whole genome physical map allows efficient and nearly unlimited isolation of genes that underpin biological mechanisms and agronomical traits. The basic methodologies (fingerprinting, assembly) for constructing such maps were established years ago and are applicable generally to any kind of genome. However, the size and features of the wheat and barley genomes require specific considerations when developing the most cost efficient strategy.

11.1 Introduction

What is the distance between two markers or two genes in a genome? Where are the genes located in the genome of interest? Are they evenly distributed or do they cluster in specific regions? How do I generate and assemble genomic sequence information in a huge genome full of repetitive DNA elements?

All of these questions relate to the structural organization of a genome; and some can be addressed by measuring genetic distances between marker alleles or phenotypes thereby building a genetic map of the region of interest (cf. Chapter 7). The genetic distance reflects the probability for the occurrence of a crossing-over during meiosis between two landmarks (molecular markers, genes) on a chromosome. This can provide a reasonable indirect measurement for the underlying physical distance given in base pairs (bp); however, since recombination is not distributed evenly along chromosomes, especially in large genome species, there is no golden rule for converting recombination frequencies into bp-distances.

N. Stein (✉)

Genebank Department, Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Gatersleben, Germany
e-mail: stein@ipk-gatersleben.de

Instead, hot spots of recombination are found in gene-rich regions while areas with low gene density or high content of constitutive heterochromatin may not show any recombination at all, even in large mapping populations derived of many thousands of gametes.

Physical mapping, on the other hand, delivers absolute distances between genomic landmarks, providing direct access to any point in a genome and to the isolation of any gene. This forms the basis for any structural and functional analysis of correlations between phenotypic and genetic information. This chapter presents the spectrum of approaches that can be used for physical mapping and illustrates current efforts for the establishment of physical maps in the Triticeae genomes.

11.2 Generating a Physical Map – Basic Principles and Methods

The development of a physical map can concern either a target locus or a complete chromosome or genome. At target loci, physical maps are established by assaying directly individual markers or genes that flank or cosegregate with the locus of interest (e.g., by selecting overlapping DNA clones or by in situ hybridization). Physical maps of chromosomes and genomes, on the other hand, are established with high-throughput generic approaches that produce reproducible, partially overlapping subunits from genomic DNA of the organism cloned into Bacterial Artificial Chromosome clones (BAC's, Shizuya et al. 1992) or Yeast Artificial Chromosome clones (YAC's, Burke et al. 1987) that can be assembled into contigs. At a subsequent stage, the physical contigs are integrated with the genetic maps using individual marker or gene information to determine the orientation of the contigs and reconstruct the chromosomes/genome (Meksem et al. 2005). Physical mapping methods can be based on ordering markers physically or by ordering overlapping clones.

11.2.1 Ordered-Marker Based Physical Mapping

The determination of physical distances at target loci is crucial for map-based gene isolation or for comparative genomic studies between closely related species. Physical distance information can be obtained through several approaches that provide varying degrees of resolution. Depending on the method, distances extending over large parts of chromosome arms (above the megabase level, microscopically visible) or distances at sub-microscopical level can be resolved. Here, we will discuss the use of cytogenetic stocks, Fluorescence In Situ Hybridization (FISH), Radiation Hybrid Mapping (RH), and Happy mapping.

11.2.1.1 Use of Cytogenetic Stocks and Chromosome-Microdissection

A cytogenetic stock is defined as a genotype that carries any kind of reproducible and maintainable chromosomal aberration that differentiates it from a normal, wild-type donor genotype. Collections comprised of a large spectrum of chromosomal aberrations offer a great opportunity for determining physical relationships between genomic landmarks. Among all plant species, the Triticeae species have the largest collection of cytogenetic stocks and these unique resources provide excellent tools for physical mapping. The allohexaploid wheat genome (*T. aestivum* L.) consists of three highly colinear and homoeologous A, B, and D genomes. This genetic redundancy allows it to tolerate full chromosome loss, partial deletion or substitution by homoeologous chromosomes from other wheat species, wild relatives, barley, or rye. The lack of physical entities of a chromosome (deletion) or the replacement of such by a homoeologous fragment from a related genome (substitution) allows determination of the presence or absence of markers or loci. Thus, landmarks that are physically close to each other on a chromosome have a higher probability to be present on the same deletion or substituted fragment. In wheat, a series of 436 segmental deletions distributed across the three homoeologous genomes has been generated (Endo and Gill 1996). For low-resolution physical mapping, a subset of 101 deletion lines (carrying several deletions) and 21 nulli-tetrasomic stocks was instrumental for the development of a wheat genome physical map comprising 7,100 expressed sequence tags (EST) (Qi et al. 2003, 2004). The characterized deletions divided the wheat genome into 189 “chromosome bins” with each providing a resolution of about 28 Mbp (Qi et al. 2003).

A complete series of whole chromosome or ditelosomic substitution lines of individual barley (*Hordeum vulgare* L.) chromosomes is available in *T. aestivum* genomic background (Shi and Endo 1997). Similar to the wheat deletion stock series, these substitution lines have been used to generate series of deletions in individual barley chromosomes (Endo 2007, 1988). Since sufficient sequence polymorphism exists between homoeologous wheat and barley genes, these wheat/barley substitution and deletion lines can be used efficiently for allocating genes to barley chromosomes. For example, substitution panels of chromosomes 5H (Fig. 11.1, Ashida et al. 2007) and 7H (Masoudi-Nejad et al. 2005) have been developed and used for physical assignment of gene-based markers. In the case of chromosome 7H, ninety deletions and/or translocations were available that provided a theoretical average resolution of >10 Mb. Chromosomal allocation of genes was achieved by single marker STS-PCR (Nasuda et al. 2005) or with high-throughput using microarray hybridizations (Bilgic et al. 2007, Cho et al. 2006). A series of barley lines carrying defined chromosomal translocations was analysed by a combined approach utilizing chromosome microdissection and PCR amplification of markers. This strategy allowed the determination of the physical position of translocation breakpoints in barley chromosomes and provided the first physical map of the barley genome (Kuenzel et al. 2000). It illustrated at genome-scale that recombination was not

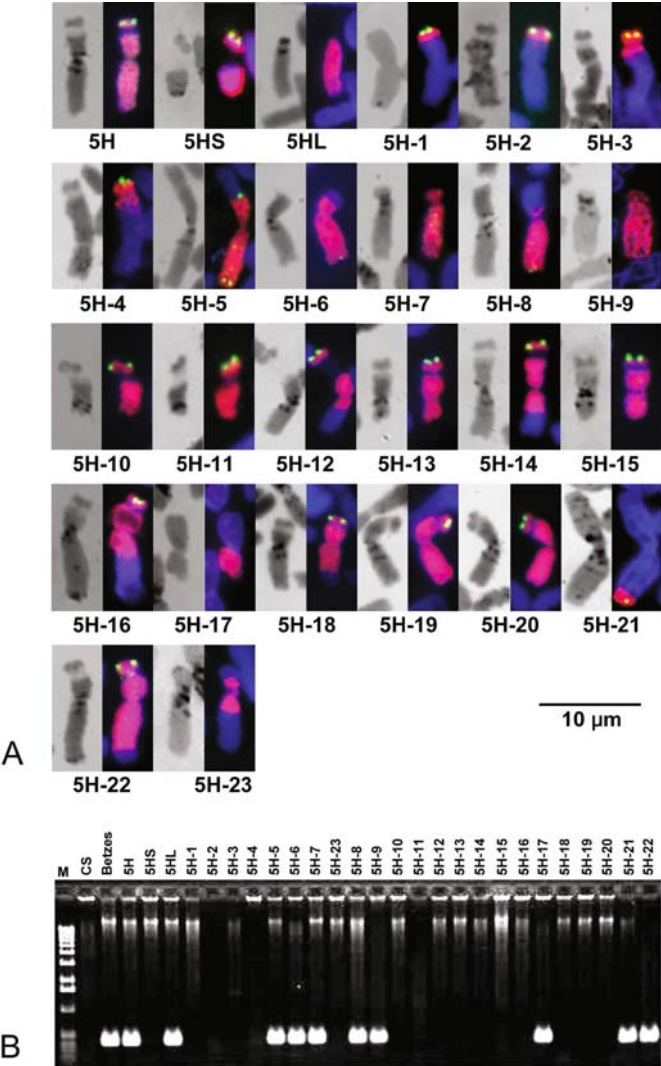


Fig. 11.1 Physical mapping using a deletion panel of barley chromosome 5H in wheat. **(A)** Barley chromosome 5H is represented as a series of deletions in the background of the *T. aestivum* genome. Barley DNA (pink) has been visualized by genomic in situ hybridization (GISH) and can be distinguished from the wheat background (blue). **(B)** STS-PCR with genomic DNA of the different deletion stocks allows for the precise allocation of a gene/ marker to a specific chromosomal fragment. The assayed marker is present on the long arm of chromosome 5H (5H-17) in a region shared by 8 overlapping deletion chromosomes. (Figure taken with permission from Ashida et al. 2007) (See Color Insert)

evenly distributed along the barley chromosomes. Indeed, the physical map indicated that, regions of high recombination frequency were located closer to telomeres while regions with low recombination frequency surrounded the centromeres of all barley chromosomes (Kuenzel et al. 2000), a feature shared also by the wheat genome (Gill 2004).

11.2.1.2 Fluorescence In Situ Hybridization (FISH)

The principle, success, and pitfalls of FISH applied in plants have been summarized in details before (Jiang and Gill 1994, 2006), thus, only specific aspects of the technology related to physical mapping in the Triticeae are addressed here.

In brief, DNA landmarks can be visualized directly on chromosomal preparations by fluorescently labeled hybridization probes. Physical distances between markers or genes can be estimated directly at the chromosomal level if multiple, differentially labeled probes are hybridized simultaneously by measuring the distance between the signals in relation to the entire length of the chromosome. Thus, the method provides, in theory, an attractive and straightforward approach for physical mapping.

While genes have been detected already by cDNA or BAC clones hybridization on barley mitotic chromosomes (Lapitan et al. 1997, Stephens et al. 2004), the method is not routinely established and cannot be applied at high-throughput in barley or wheat due to issues affecting reproducibility. Primarily, this is a result of limitations in the accessibility of chromosomes in surface spread or squash preparations of chromosomes, and the difficulty of developing probes (probes <10 kb are difficult to detect whereas probes >10 kb increase the risk of bearing repetitive DNA leading to spurious signal noise). Furthermore, physical resolution which is in the range of 5–10 Mb if working with mitotic chromosomes is limited also (Pedersen and Linde-Laursen 1995). Improved resolution can be achieved by utilizing less condensed targets (Cheng et al. 2002, Valarik et al. 2004). In plants where less condensed meiotic prophase (i.e. pachytene) chromosomes are accessible easily and can be addressed specifically in a reproducible manner, FISH allows the determination of physical distances at a resolution of less than 100 kb. In maize, this approach facilitated the determination of physical distances between markers on chromosome 9 which is easy to recognize at pachytene (Sadder et al. 2002, Wang et al. 2006). As shown in rice, the possibility to visualize overlapping BAC clones forming Minimal Tiling Paths (MTPs) by FISH permits the validation of physical maps or the determination of gap sizes that remain in BAC contig maps (Cheng et al. 2005, 2002). In barley and wheat, pachytene chromosomes have not been used for FISH physical mapping, possibly due to difficulties in the specific identification and separation of individual chromosomes at this stage. An alternative approach is to work on purified individual chromosomes as Triticeae chromosomes can be isolated and purified by flow-cytometry (Doležel et al. 2005; and Chapter 10) thereby providing a superior target for FISH. When artificially stretched on the

surface of microscopic slides, the resolution obtained is even superior to that of pachytene chromosomes (Valarik et al. 2004). Higher resolution, i.e. at a level of <10 kb, can be achieved if working on stretched DNA fibers; however, a direct correlation of the results to any chromosome morphology that is characteristic for linkage group identity is no longer feasible (Cheng et al. 2002, Doležel et al. 2005). In contrast to the accuracy that is potentially achievable, FISH-based methods lack the throughput and resolution needed for the construction of sequence ready physical maps in the Triticeae. Therefore, FISH is used in conjunction with generic physical mapping of genomes and serves only as a concomitant and case-by-case verification tool to prove physical map consistency and to sample gap sizes remaining in contig maps.

11.2.1.3 Radiation Hybrid Mapping (RH) – HAPPY Mapping

Markers or genes with close physical linkage in the genome will be identified on the same fragment of randomly sheared genomic DNA at a higher probability than those separated by larger physical distances. To exploit this principle, radiation hybrid mapping (RH) and HAPPY (HAP = haploid cells, PY = polymerase chain reaction) mapping were designed (reviewed in Waugh et al. 2002).

For RH mapping, hybrid cell lines are formed through fusion of a lethally irradiated cell line (donor) of one species with a normal viable cell line (acceptor) of another species (Cox 1992, Goss and Harris 1975). The donor cell line receives a lethal radiation dose before fusion that leads to numerous random chromosome breaks. After cell fusion, these donor chromosome segments integrate at random into the acceptor cell genome and, thus, become stabilized and remain integrated subsequently. Each individual hybrid cell gives rise to a clone representing a random sub-portion of the donor genome. Since donor and acceptor genomes originate from different species, specific markers for the donor genome can be assayed on the DNA stocks of a comprehensive RH mapping panel. The frequency by which linked probes are transmitted together to independent clones allows the estimation of the linear order of genes or markers using the physical distance between genes and by the radiation dose applied to the donor cell line. RH panels are relatively easy to develop in animal systems and have been employed in numerous physical mapping and sequencing projects of mammalian genomes. This culminated in a proposition that every low-pass genome survey sequencing project should be accompanied with a dense gene-based RH map (Hitte et al. 2005).

So far, radiation hybrid mapping has played only a limited role in physical mapping of plant genomes. Since cell fusion in plants requires removal of the cell wall, efficient systems to generate protoplasts from the plant species of interest must be established. Further, fusion products need a high capacity for

regeneration into callus to ensure maintenance of the resource in tissue culture. Such a system has been established once in plants by fusing transgenic bialaphos resistant barley to tobacco protoplasts (Wardrop et al. 2002). Using transgenic barley as a donor allowed the selection of fusion products on a bialaphos containing medium to increase the recovery rate of clones carrying barley DNA. Pilot screenings generally demonstrated the feasibility of the approach but scarcity of the DNA resource that was obtained from the regenerating callus limited its potential for large scale screening. In combination with full genome amplification, Wardrop et al. (2004) have shown that this limitation possibly can be overcome. However, broad application and usefulness of the resource remains to be proven.

Another approach, perhaps more sustainable, was followed for maize and wheat. Single chromosome addition (single maize chromosomes in a hexaploid oat background) or substitution lines (*Triticum aestivum* D-genome chromosomes in tetraploid durum wheat background), respectively, were exposed to high radiation doses and were backcrossed to a male-sterile isogenic line resulting in genotype panels that carried random fragments of the addition/substitution chromosome. These collections, which resemble RH panels, allowed physical mapping at a resolution of 0.2–1 Mb (Kalavacharla et al. 2006, Riera-Lizarazu et al. 2000). Since these RH stocks can be maintained and distributed as seeds, they even have advantages over classical RH panels. Genetic stocks are in place to manipulate all maize (Kynast et al. 2004) and wheat D-genome (Joppa 1993, Joppa and Williams 1988) chromosomes in similar manner and, therefore, RH mapping should be considered as a true option for physical mapping in these genomes. This will be particularly useful for the integration of physical contigs that originate from the low or no recombinogenic proximal regions of the wheat and barley chromosomes. A RH panel has been established and is currently under evaluation (C. Feuillet and S. Kianian pers. comm.) to support the construction of the physical map of chromosome 3B (see below).

HAPPY mapping is an *in vitro* alternative to RH mapping that relies on the same mapping principle (Dear and Cook 1989, Waugh et al. 2002). DNA of the genome of interest is sheared mechanically and then dispensed into aliquots that represent less than one genome equivalent – too minute for assaying directly markers or genes. However, methods of “whole genome amplification” (Dean et al. 2002, Lage et al. 2003, Telenius et al. 1992, Zhang et al. 1992) allow amplifying the contained DNA into quantities that then allow for thousands of individual screens with the same amplified DNA stocks. To date, the only example for the application of HAPPY mapping in plants was reported in *A. thaliana* (Thangavelu et al. 2003) where a resolution at below 50 kb was achieved. However, in the same study, the authors stated that the potential of HAPPY mapping for >2.5 Gb genomes, such as those of the Triticeae, will be limited likely to restricted gene-rich areas and will not allow the development of a genome-scale HAPPY map.

11.2.2 Ordered-Clone Based Physical Mapping

Building a whole genome physical map can be performed with a statistical non-biased approach. Based on the assumption that in a redundant genomic library every part of a given genome is represented by an overlapping set of clone inserts, it is theoretically possible to reconstruct the genome by ordering contiguously overlapping clones (forming a so-called “contig”). Initially, the genome of interest is cloned into a large insert library, today, usually a BAC library (Fig. 11.2A). Large numbers of clones, typically representing an equivalent of about 15-fold genome coverage or higher, need to be analyzed to reach sufficient genome representation in contigs and contig length. However, depending on the complexity of the genome, this approach will not lead directly to single contigs for each chromosomes but rather to a few thousand independent contigs (Meyers et al. 2004). For example, in maize (~3 Gb haploid genome size), fingerprinting of 15- to 18-fold genome coverage by either agarose-gel based fingerprinting or by HICF resulted in 4,518 and 1,500 FPC contigs, respectively (Wei et al. 2007). A redundant set of contiguously overlapping clones will be identified out of which a non-redundant set of clones, a so-called “Minimal Tiling Path” (MTP) will be selected for subsequent applications (Fig. 11.2B). For efficient utilization of the data, the MTP must be ordered along a high-density genetic map (Fig. 11.2C) that is achieved by probing the genomic BAC library with large numbers of genetically mapped markers (gene-derived or of anonymous origin), thus connecting clone addresses (as landmarks in the

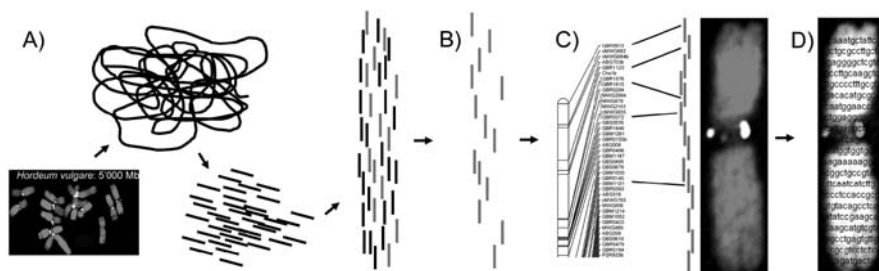


Fig. 11.2 Construction of a physical map using ordered genomic clones. (A) Genomic DNA is fragmented and cloned into a vector (fosmid, BAC, YAC), leading to a large insert genomic library. After characterization of all the clones by fingerprinting [e.g. High Information Content Fingerprinting (HICF)] and comparison of the fingerprints, overlapping clones are assembled into contigs. This will result into a redundant set of overlapping clones that reproduce the complete genome. (B) A non-redundant set of clones representing a Minimal Tiling Path (MTP) can then be selected for any further work. (C) To allocate the contig map to a chromosome or a linkage group, parts of the contig have to be anchored to maps that provide information of the relative order of markers i.e. genetic maps or RH maps. Typically, markers located on a genetic map are probed against the large-insert genomic library used for building the physical contig map and positive clones respectively their corresponding contigs are assigned to a position on the chromosome/linkage group. (D) The ordered MTP provides a template for a clone-by-clone genome sequencing

contig/MTP) with marker IDs. Such a genetically anchored map can be used immediately for positional cloning of genes. Further, a high-quality verified and genetically anchored MTP is also the template for generating a high-quality reference sequence of a genome (Fig. 11.2D).

Analyzing large numbers of genomic clones requires a generic approach applicable to high-throughput analysis. The first attempts, developed in the framework of the *Caenorhabditis elegans* genome project, were based on enzymatic restriction of cosmid clone inserts, radio-labelling of restriction fragments, gel-electrophoretic fragment separation, and subsequent fragment visualization *via* autoradiography (Coulson et al. 1986). Rapidly, more efficient methods were established that were based only on the identification of conserved restriction fragment patterns of the large-insert genomic clones obtained after agarose gel-electrophoresis – a strategy that was applied efficiently during the yeast and the human genome projects (Marra et al. 1997, Olson et al. 1986). Since then, genome-scale ordered-clone physical mapping has developed into high-throughput analysis of large-insert clones. Fingerprints are obtained either after restriction digestion with a single restriction enzyme (Marra et al. 1997) (Fig. 11.3) or with a cocktail of different specific endo-nucleases (Wong et al. 1997) (Fig. 11.4A). Excellent reviews of the complete panel of high-throughput fingerprinting techniques and their application for physical map construction in plants have recently been published (Meyers et al. 2004, Ren et al. 2005, Wu et al. 2005) and thus will not be presented here in more detail.

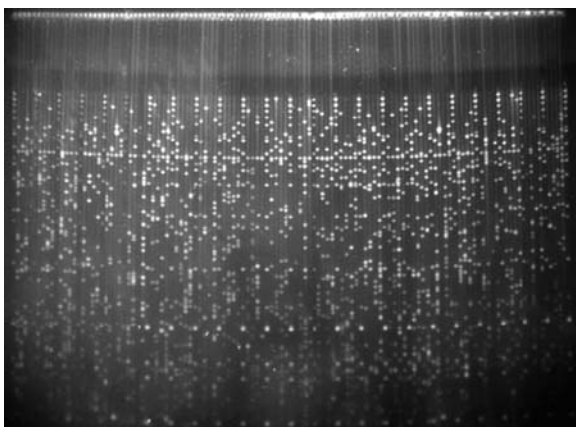


Fig. 11.3 High-throughput BAC clone fingerprinting I. BAC clone DNA cleaved by the restriction endonuclease *Hind* III is separated by electrophoresis on a highly standardized, fluorescently stained agarose gel. Such gel image is subsequently interpreted with the help of image analysis software (e.g. Image, Sulston et al. 1989, 1988) to record fingerprinting pattern of the individual clones. In order to allow reliable determination of fragment sizes, a size standard is loaded on every fifth lane of the gel. (Image kindly provided by Dr. Rod Wing, AGI, Arizona, USA)

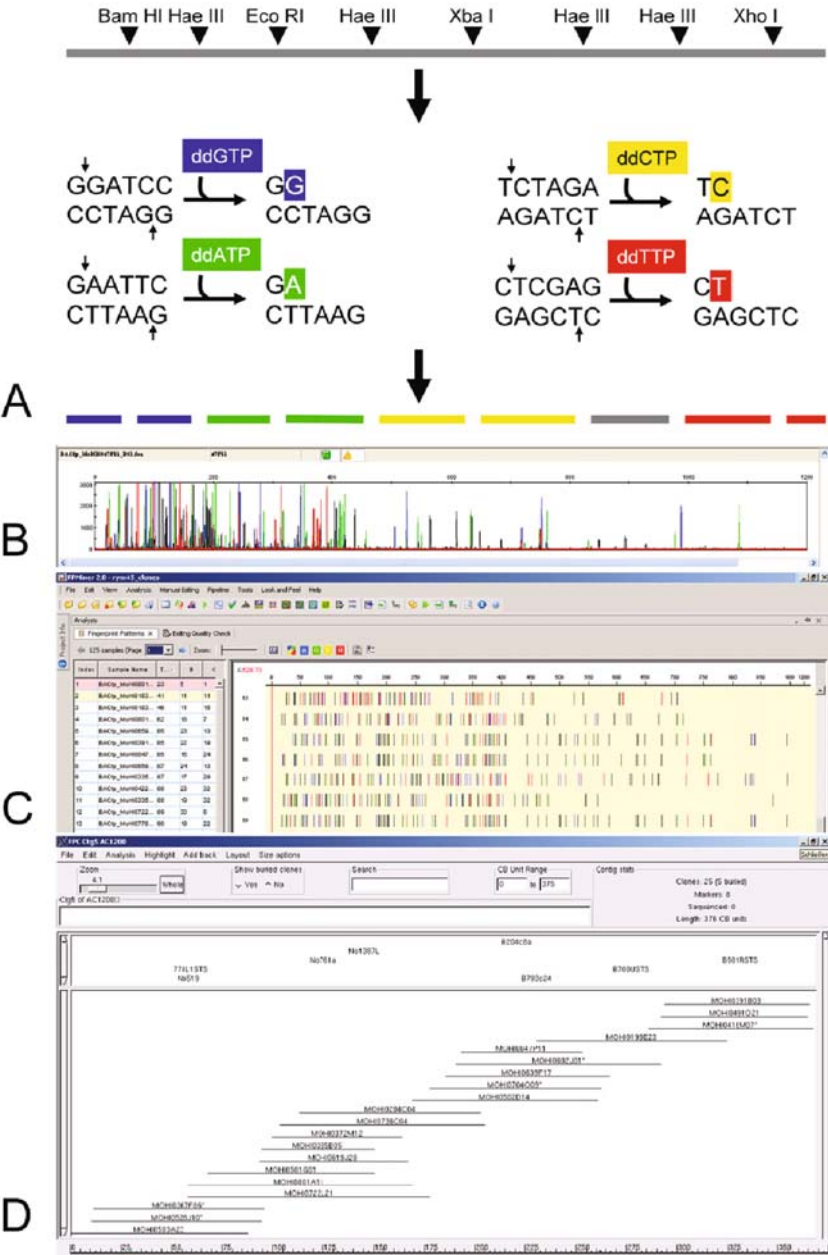


Fig. 11.4 High-throughput BAC clone fingerprinting II. (A) Schematic representation of multi-color HICF (Luo et al. 2003a). Insert DNA is subjected to multi-enzyme restriction digestion: the four endonucleases *Bam* HI, *Eco* RI, *Xba* I and *Xho* I cut at their respective hexamer recognition site, all producing fragment ends that differ at the first base of the protruding DNA strands. This information is utilized in a subsequent strand-synthesis reaction after adding differently fluorescence-labeled dideoxy-nucleotides (SnapShot, Applied Biosystems)

High-throughput fingerprinting became possible by the availability of image analysis algorithms that allow the acquisition of fingerprinting raw data information (Image: Sulston et al. 1989, 1988) and computational tools for automated physical map assembly based on fingerprint information (fingerprint contigs FPC: Soderlund et al. 2000, 1997). In the simplest case, BAC DNA is cut by a single enzyme (e.g., *HindIII*) and the resulting fragments are separated by highly standardized agarose-electrophoresis and visualized by subsequent staining in ethidium-bromide or SYBRgreen (Fig. 11.3). The size of the fragments obtained from a single clone is determined and the fingerprint pattern is recorded. A combination of fluorescent labeling of DNA fragments and resolving fluorescent fingerprint patterns by the use of automated capillary gel-electrophoresis advanced the approach to high-information content fingerprinting (HICF) (Ding et al. 2001, 1999, Luo et al. 2003a). Similar to automated fluorescent Sanger-Sequencing, the fingerprinting results are recorded as electropherograms (Fig. 11.4) and can be interpreted by the FPC software (Soderlund et al. 1997) similar to non-fluorescent single-enzyme based fingerprints.

11.2.2.1 Chromosome Walking

Positional or map-based cloning relies on the principle of chromosome walking that is the process of generating a local ordered-clone physical map around the gene of interest. The approach employs and takes advantage of the same resources that are required to build an ordered-clone based physical map of a genome (see above) but at a reduced scale. The principle of chromosome walking has been described in details before (Jander et al. 2002, Leyser and Chang 1996, Peters et al. 2003), even with emphasis on grass genomes including the Triticeae (Stein and Graner 2004) and, thus, this aspect of physical mapping will not be described further here. Recent advances of map-based cloning in the Triticeae genomes are reviewed in Chapter 12 of this book.

11.2.3 Optical Mapping

Perhaps, the most sophisticated approach for accumulating physical mapping information of a genome is optical mapping (Schwartz et al. 1993). In this

◀ **Fig. 11.4** (continued) leading to a mixture of accordingly labeled DNA fragments. A fifth endonuclease with a four bp-recognition site, *Hae III* is added to the reactions to reduce the overall average fragment length. **(B)** Separation of the reaction on a capillary sequencer results in characteristic so-called electropherograms (multicolor histograms) with every peak standing for a specifically labeled DNA fragment. **(C)** The individual peaks are converted into bands using Genoprofiler (You et al. 2007) to produce a virtual gel-image. **(D)** The fragment patterns of different genomic clones are compared using the software FPC (Soderlund et al. 1997) and clones can be assembled into contigs based on the statistical analysis of the level of similarity between their fragments (See Color Insert)

method, pure DNA molecules (i.e., linearized BACs, YACs or sheared Mb-sized genomic DNA) of a genome become surface-stretched and attached to a microscopic glass slide (Dimalanta et al. 2004, Jing et al. 1998). After staining of the DNA molecules, the specimen can be visualized by light microscopy and digital imaging devices. Subsequently, DNA is treated with a rare-cutting restriction endonuclease that introduces gaps into the molecule. The order of the differently sized restriction fragments of one molecule is recorded and pair wise comparisons are performed with patterns of other molecules present in the same sample. This approach was validated and applied successfully to a number of microbial genomes resulting in detailed optical restriction maps (Aston et al. 1999). Only recently, the approach was applied successfully to rice demonstrating the feasibility of optical mapping in higher eukaryotic organisms with larger genomes (Zhou et al. 2007). The general applicability of optical mapping remains limited, however. It can be employed best in helping to verify whole genome sequence assemblies through comparisons of the *in silico* restriction map of a genome sequence assembly with the whole genome optical map obtained *in situ* by the same endonuclease. In principle, to take full advantage of the resource, whole genome optical maps can be integrated to an ordered-clone physical map by assaying short tiling paths of 10–30 BAC clones with the same approach and with the same enzyme used for whole genome optical mapping and the resulting restriction pattern of the BAC pools is integrated then into the genome optical map (Aston et al. 1999).

Although the approach was successfully applied to a small plant genome, the general applicability and cost efficiency of optical mapping to >2.5 Gb genomes such as those of the Triticeae remains an open question. Based on the available literature, optical mapping has been applied exclusively in the laboratory where it was originally developed. This may be due to the lack of necessary technical infrastructure in other labs or, perhaps, this provides an indication of the difficulties that exist in transferring the skills to other laboratories. Thus, the usefulness of optical mapping for *de novo* physical mapping of whole genomes remains to be demonstrated.

11.3 Physical Maps of Triticeae Genomes

The construction of physical maps of the Triticeae (wheat, barley, rye) genomes is an essential prerequisite for accelerated isolation of agronomically important genes and for genome sequencing. Despite individual successes obtained after years of laborious work (Stein and Graner 2004) (see also Chapter 12), map based cloning in wheat and barley has remained inefficient with respect to the need for a better understanding of the basis of agronomically important traits and for accelerated improvement of these essential crop species. Genome size has a strong impact on the process of building a genome-wide physical map. The smallest Triticeae crop genomes are in the size range

of about 5–6 Gbp (e.g., barley, Flavell et al. 1974) which is almost twice the size of the largest eukaryotic genomes, such as human (3.2 Gbp, The International Human Genome Mapping 2001) and maize (2.3 Gbp, Rayburn et al. 1993), for which whole genome BAC fingerprinting has been achieved (The International Human Genome Mapping 2001, Wei et al. 2007). Simulations based on algorithms developed for calculating assemblies of a fingerprinting or shotgun sequencing project (Lander and Waterman 1988, Wendl and Waterston 2002) show that with an average clone insert size of 120 kb, fingerprinting coverage of 15-fold haploid genome equivalents should be performed to obtain a few thousand independent contigs representing these genomes. In a diploid genome of 5 Gbp, this represents fingerprinting more than 600,000 BAC clones; whereas, for polyploid Triticeae species like pasta wheat (*T. durum*, 4n, AB) and bread wheat (*T. aestivum*, 6n, ABD), this number increases by a factor of two to three (i.e., about 2 million clones). Moreover, HICF in a polyploid genome can be complicated further by sequence conservation between the homoeologous genomes which may lead to false co-assemblies of large-insert clones that originate from highly conserved homoeologous loci (e.g., gene rich regions). In the past few years, wheat and barley scientists have addressed the challenges by building resources and adapting methods that have allowed the development of physical mapping projects for three Triticeae crop species (Table 11.1).

Table 11.1 Summary of ongoing projects on physical mapping Triticeae genomes

Species	Type of map	Clone number	coverage	reference
<i>Ae. tauschii</i>	Whole genome	270,000	6.6	Luo et al. 2003b
<i>H. vulgare</i>	Gene-space	65,000	—	Madishetty et al. 2007
<i>H. vulgare</i>	Whole genome	200,000	4.0	Own unpublished data ¹
<i>T. aestivum</i>	Chromosome 3B	68,000	8.0	Paux et al. 2008
<i>T. aestivum</i>	Chromosome 3AS	56,000	12	Unpublished data ²
<i>T. aestivum</i>	Chromosome 1A	—	—	Unpublished data ²
<i>T. aestivum</i>	Chromosome 1B	—	—	Unpublished data ²
<i>T. aestivum</i>	Chromosome 3D	—	—	Unpublished data ²

¹ project in progress, final target is 12–15-fold coverage = >500,000 clones HICF, ² USDA–NRI project (personal communication, Prof. B. Gill, Kansas State University, USA). Progress of the *T. aestivum* and *Ae. tauschii* physical mapping projects can be followed on the International Wheat Genome Sequencing Consortium (IWGSC) website at www.wheatgenome.org.

11.3.1 Physical Maps of Diploid Triticeae Genomes

11.3.1.1 *Aegilops Tauschii*

The first project for the construction of a whole genome physical map in a Triticeae species was initiated in the wild diploid wheat, *Ae. tauschii*, the

ancestral donor species of the D genome of *T. aestivum*. After establishing a five enzyme multicolor HICF protocol based on the SNaPshot (Applied Biosystems) labeling technology (Luo et al. 2003a) as well as developing new software for editing fingerprinting information (Genoprofiler, You et al. 2007), an assembly (Sulston score = e-25, tolerance = 0.4 bp) of nearly 270,000 BAC clones was performed that resulted in about 13,000 contigs (Luo et al. 2003b, updated in 2004 at <http://wheatdb.ucdavis.edu:8080/wheatdb/>). With an average insert size of 110 kb and a genome size of 4500 Mbp for *Ae. tauschii*, this represents a haploid genome coverage of about 6.6 fold. The BAC libraries used for fingerprinting were probed with 508 RFLP and 219 EST markers for anchoring contigs to the genetic map. The assembly and project information is available at <http://wheatdb.ucdavis.edu:8080/wheatdb/>. A proposal has been approved recently by the National Science Foundation, USA, to complete the above mentioned project and reach higher genome coverage for the *Ae. tauschii* physical map so that it can be used as a framework for the physical map of the *T. aestivum* D-genome (<http://www.wheat-genome.org/projects.html>, “US-NSF PGRP 2006, Physical Mapping of the wheat D-Genome – *Ae. tauschii* and 1D, 4D, and 6D of hexaploid wheat”, PI = Jan Dvorak, UC Davis, USA).

11.3.1.2 Barley (*Hordeum vulgare*)

The only diploid Triticeae crop species for which physical mapping of the whole genome is underway is barley; initially focused exclusively to fingerprinting of gene containing regions (NSF #0321756, <http://phymap.ucdavis.edu:8080/barley/>). A BAC library representing 6.3-fold haploid genome coverage (Yu et al. 2000) was screened by hybridization with 12,600 overgo probes resulting in the identification of about 65,000 BAC clone addresses (Madishetty et al. 2007). Based on the observation that genes are not distributed randomly in the barley genome (Varshney et al. 2006), the hypothesis was that the presence of a gene on a BAC clone increased the potential for the presence of a second gene on the same BAC. Thus, by following this approach a majority of gene containing BAC clones should be identified. All positive clones were analyzed by HICF following the method of (Luo et al. 2003a) and about 10,000 contigs were obtained (status June 2008, <http://phymap.ucdavis.edu:8080/barley/>). To extend this effort further, a project to develop a whole genome physical map was initiated in June 2006 in the framework of the International Barley Genome Sequencing Consortium (IBSC, <http://barleygenome.org>) with the goal of reaching 12- to 15-fold haploid genome coverage of the barley genome by HICF. This is expected to deliver about 7000 contigs by the end of 2009. To date, more than 200,000 clones have been fingerprinted for a 4-fold haploid genome coverage (own unpublished data, project progress can be followed at the IBSC website <http://barleygenome.org>).

11.3.2 Physical Maps of Polyploid Triticeae Genomes

11.3.2.1 Bread wheat (*Triticum aestivum*)

Bread wheat has one of the largest and most complex crop plant genomes. The huge size (17.33 Gbp; Bennett and Smith 1976), the presence of three highly related and collinear, homoeologous genomes (A, B, D) and the high sensitivity of the HICF method raised controversial discussions about the cost efficiency and feasibility of building a physical map of the complete wheat genome by HICF in an international effort. Thus, the International Wheat Genome Sequencing Consortium (IWGSC, <http://www.wheatgenome.org/>) decided to evaluate an alternative strategy based on the construction of individual chromosome or chromosome arm physical maps. Recent advances in flow cytometry combined with the availability of appropriate wheat cytogenetic stocks allowed the wheat genome to be dissected into individual chromosomes and chromosome arms (Doležel et al. 2007) (see Chapter 10). In this way, the bread wheat genome can be divided into segments as small as 1.3% of the total genome, i.e. about 220 Mbp or approximately half of the size of the fully sequenced rice genome. With this technique, it is possible to isolate high-molecular weight DNA in quantity and quality that is suitable for the construction of chromosome-specific BAC libraries (Doležel et al. 2007, Safar et al. 2004). A first pilot project on wheat chromosome 3B, the largest wheat chromosome (~1 Gb, >2 times larger than the rice genome) has demonstrated recently the feasibility of the approach. Nearly 68,000 clones of a chromosome 3B specific BAC library (Safar et al. 2004) with an average insert size of 120 kbp were fingerprinted using the HICF method developed by (Luo et al. 2003a) and about 1,000 contigs representing 82% of the chromosome were assembled (Paux et al. 2008). Anchoring of the contigs has been performed using SSR and EST markers known to be located on chromosome 3B as well as SSR and ISBP markers (Paux et al. 2006) derived from BAC end sequences. To date, 66% of the map has been anchored to genetic and cytogenetic maps. Using the cytogenetic stocks (ditelosomic and isochromosomic lines), all 21 individual chromosome or chromosome arms can be sorted and used for BAC library construction and physical mapping. Projects are underway already for chromosome 3AS and 3AL as well as for chromosomes 1A, 1B, and 3D (Table 11.1; Feuillet and Eversole 2008) and several international members of the IWGSC are preparing grant applications for the rest of the chromosomes. Progress towards physical mapping of the hexaploid wheat genome can be followed at <http://www.wheatgenome.org/>.

11.4 Conclusion

Already, physical mapping has a long tradition in the Triticeae through the use of cytogenetic stocks. Only recently, the convergent development of genomic resources (BAC libraries, dense genetic maps, marker collections) and efficient

methods (high-throughput fingerprinting, chromosome sorting) have opened a perspective to tackle the most economically important and most technically challenging Triticeae genomes for genome-wide physical map construction. These physical maps will accelerate greatly gene isolation, and crop improvement through marker assisted selection and genetic engineering. In times where a revolution is taking place in the field of sequencing technologies, such physical maps also represent the foundation for systematic sequencing of the Triticeae genomes (Chapter 24).

References

- Ashida T, Nasuda S, Sato K, et al. (2007) Dissection of barley chromosome 5 H in common wheat. *Genes Genet Syst* 82: 123–133
- Aston C, Mishra B, Schwartz D C (1999) Optical mapping and its potential for large-scale sequencing projects. *Trends Biotechnol* 17: 297–302
- Bennett M D, Smith J B (1976) Nuclear DNA amounts in angiosperms. *Philos Trans R Soc Lond B Biol Sci* 274: 227–274
- Bilgic H, Cho S, Garvin D F et al. (2007) Mapping barley genes to chromosome arms by transcript profiling of wheat–barley ditelosomic chromosome addition lines. *Genome* 50: 898–906
- Burke D T, Carle G F, Olson M V (1987) Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. *Science* 236: 806–812
- Cheng C H, Chung M C, Liu S M, et al. (2005) A fine physical map of the rice chromosome 5. *Mol Genet Genom* 274: 337–345
- Cheng Z, Buell C, Wing R, et al. (2002) Resolution of fluorescence in-situ hybridization mapping on rice mitotic prometaphase chromosomes, meiotic pachytene chromosomes and extended DNA fibers. *Chromosome Res* 10: 379–387
- Cho S, Garvin D F, Muehlbauer G J (2006) Transcriptome analysis and physical mapping of barley genes in wheat–barley chromosome addition lines. *Genetics* 172: 1277–1285
- Coulson A, Sulston J, Brenner S, et al. (1986) Toward a physical map of the genome of the nematode *Caenorhabditis elegans*. *Proc Natl Acad Sci USA* 83: 7821–7825
- Cox D R (1992) Radiation hybrid mapping. *Cytogenet Cell Genet* 59: 80–81
- Dean F B, Hosono S, Fang L, et al. (2002) Comprehensive human genome amplification using multiple displacement amplification. *Proc Natl Acad Sci* 99: 5261–5266
- Dear P H, Cook P R (1989) HAPPY mapping: A proposal for linkage mapping the human genome. *Nucleic Acids Res* 17: 6795–6807
- Dimalanta E T, Lim A, Runnheim R, et al. (2004) A microfluidic system for large DNA molecule arrays. *Anal Chem* 76: 5293–5301
- Ding Y, Johnson M D, Chen W Q et al. (2001) Five-color-based high-information-content fingerprinting of bacterial artificial chromosome clones using type IIS restriction endonucleases. *Genomics* 74: 142–154
- Ding Y, Johnson M D, Colayco R, et al. (1999) Contig assembly of bacterial artificial chromosome clones through multiplexed fluorescence-labeled fingerprinting. *Genomics* 56: 237–246
- Doležel J, Kubaláková M, Bartoš J, et al. (2005) Chromosome flow sorting and physical mapping. In: Meksem K, Kahl G (eds.), *The handbook of plant genome mapping. Genetic and physical mapping*. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, 151–171
- Doležel J, Kubaláková M, Paux E, et al. (2007) Chromosome-based genomics in the cereals. *Chromosome Res* 15: 51–66

- Endo T (2007) The gametocidal chromosome as a tool for chromosome manipulation in wheat. *Chromosome Res* 15: 67–75
- Endo T R (1988) Induction of chromosomal structural changes by a chromosome of *Aegilops cylindrica* L. In common wheat. *J Hered* 79: 366–370
- Endo T R, Gill B S (1996) The deletion stocks of common wheat. *J Hered* 87: 295–307
- Feuillet C, Eversole K (2008) Physical mapping of the wheat genome: A coordinated effort to lay the foundation for genome sequencing and develop tools for breeders. *Isr J Plant Sci* 55: 307–313
- Flavell R B, Bennett M D, Smith J B, et al. (1974) Genome size and the proportion of repeated nucleotide sequence DNA in plants. *Biochem Genet* 12: 257–269
- Gill K S (2004) Gene distribution in cereal genomes. In: Gupta PK, Varshney RK (eds.), *Cereal genomics*. Kluwer Academic Publishers, Dordrecht, Boston, London, pp 361–384
- Goss S J, Harris H (1975) New method for mapping genes in human chromosomes. *Nature* 255: 680–684
- Hitte C, Madeoy J, Kirkness E F, et al. (2005) Facilitating genome navigation: Survey sequencing and dense radiation-hybrid gene mapping. *Nat Rev Genet* 6: 643–648
- Jander G, Norris S R, Rounsley S D, et al. (2002) Arabidopsis map-based cloning in the post-genome era. *Plant Physiol* 129: 440–450
- Jiang J, Gill B S (1994) Nonisotopic in situ hybridization and plant genome mapping: The first 10 years. *Genome* 37: 717–725
- Jiang J, Gill B S (2006) Current status and the future of fluorescence in situ hybridization (fish) in plant genome research. *Genome* 49: 1057–1068
- Jing J, Reed J, Huang J, et al. (1998) Automated high resolution optical mapping using arrayed, fluid-fixed DNA molecules. *Proc Natl Acad Sci USA* 95: 8046–8051
- Joppa L (1993) Chromosome engineering in tetraploid wheat. *Crop Sci* 33: 908–913
- Joppa L, Williams N (1988) Langdon durum disomic substitution lines and aneuploid analysis in tetraploid wheat. *Genome* 30: 222–228
- Kalavacharla V, Hossain K, Gu Y, et al. (2006) High-resolution radiation hybrid map of wheat chromosome 1D. *Genetics* 173: 1089–1099
- Kuenzel G, Korzun L, Meister A (2000) Cytologically integrated physical restriction fragment length polymorphism maps for the barley genome based on translocation breakpoints. *Genetics* 154: 397–412
- Kynast R G, Okagaki R J, Galatowitsch M W, et al. (2004) Dissecting the maize genome by using chromosome addition and radiation hybrid lines. *Proc Natl Acad Sci USA* 101: 9921–9926
- Lage J M, Leamon J H, Pejovic T, et al. (2003) Whole genome analysis of genetic alterations in small DNA samples using hyperbranched strand displacement amplification and array-cgh. *Genome Res* 13: 294–307
- Lander E S, Waterman M S (1988) Genomic mapping by fingerprinting random clones: A mathematical analysis. *Genomics* 2: 231–239
- Lapitan N L V, Brown S E, Kennard W, et al. (1997) Fish physical mapping with barley BAC clones. *Plant J* 11: 149–156
- Leyser O, Chang C (1996) Chromosome walking. In: Foster GD, Twell D (eds.), *Plant gene isolation*. John Wiley and Sons, Ltd, Chichester, pp 248–271
- Luo M-C, Thomas C, You F M, et al. (2003a) High-throughput fingerprinting of bacterial artificial chromosomes using the snapshot labeling kit and sizing of restriction fragments by capillary electrophoresis. *Genomics* 82: 378–389
- Luo M C, Thomas C S, Deal K R, et al. (2003b) Construction of contigs of *Aegilops tauschii* genomic DNA fragments cloned in BAC and BIBAC vectors. In: *10th International Wheat genetics Symposium*, September 1–6, 2003, Paestum, Italy, pp 293–296
- Madishetty K, Condamine P, Svensson J T, et al. (2007) An improved method to identify bac clones using pooled overgos. *Nucleic Acids Res* 35: e5
- Marra M A, Kucaba T A, Dietrich N L, et al. (1997) High throughput fingerprint analysis of large-insert clones. *Genome Res* 7: 1072–1084

- Masoudi-Nejad A, Nasuda S, Bihoreau M-T, et al. (2005) An alternative to radiation hybrid mapping for large-scale genome analysis in barley. *Mol Genet Genom* 274: 589–594
- Meksem K, Ishihara H, Jesse T (2005) Integration of physical and genetic maps. In: Meksem K, Kahl G (eds.), *The handbook of plant genome mapping. Genetic and physical mapping*. WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim, 215–232
- Meyers B C, Scalabrin S, Morgante M (2004) Mapping and sequencing complex genomes: Let's get physical! *Nat Rev Genet* 5: 578–588
- Nasuda S, Kikkawa Y, Ashida T, et al. (2005) Chromosomal assignment and deletion mapping of barley est markers. *Genes Genet Syst* 80: 357–366
- Olson M V, Dutchik J E, Graham M Y et al. (1986) Random-clone strategy for genomic restriction mapping in yeast. *Proc Natl Acad Sci USA* 83: 7826–7830
- Paux E, Roger D, Badaeva E, et al. (2006) Characterizing the composition and evolution of homoeologous genomes in hexaploid wheat through bac-end sequencing on chromosome 3B. *Plant J* 48: 463–474
- Paux E, Sourdille P, Salse J, Saintenac C, Choulet F, Leroy P, Korol A, Michalak M, Kianian S, Spielmeier W et al. (2008) A Physical Map of the 1-Gigabase Bread Wheat Chromosome 3B. *Science* 322: 101–104
- Pedersen C, Linde-Laursen I (1995) The relationship between physical and genetic distances at the *hor1* and *hor2* loci of barley estimated by two-colour fluorescent in situ hybridization. *Theor Appl Genet* 91: 941–946
- Peters J L, Cnudde F, Gerats T (2003) Forward genetics and map-based cloning approaches. *Trends Plant Sci* 8: 484–491
- Qi L, Echalié B, Friebe B, et al. (2003) Molecular characterization of a set of wheat deletion stocks for use in chromosome bin mapping of ESTs. *Funct. Integr. Genomics* 3: 39–55
- Qi L L, Echalié B, Chao S, et al. (2004) A chromosome bin map of 16,000 expressed sequence tag loci and distribution of genes among the three genomes of polyploid wheat. *Genetics* 168: 701–712
- Rayburn A, Biradar D, Bullock D, et al. (1993) Nuclear DNA content in F1 hybrids of maize. *Heredity* 70: 294–300
- Ren C, Xu Z, Sun S, et al. (2005) Genomic DNA libraries and physical mapping. In: Meksem K, Kahl G (eds.), *The handbook of plant genome mapping. Genetic and physical mapping*. WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim, 173–214
- Riera-Lizarazu O, Vales M I, Ananiev E V et al. (2000) Production and characterization of maize chromosome 9 radiation hybrids derived from an oat-maize addition line. *Genetics* 156: 327–339
- Sadder M T, Weber G (2002) Comparison between genetic and physical maps in *Zea mays* L. of molecular markers linked to resistance against *Diatraea* spp. *Theor Appl Genet* 104: 908–915
- Safar J, Bartos J, Janda J, et al. (2004) Dissecting large and complex genomes: Flow sorting and BAC cloning of individual chromosomes from bread wheat. *Plant J* 39: 960–968
- Schwartz D C, Li X, Hernandez L, I et al. (1993) Ordered restriction maps of *saccharomyces cerevisiae* chromosomes constructed by optical mapping. *Science* 262: 110–114
- Shi F, Endo T (1997) Production of wheat-barley disomic addition lines possessing an *aegilops cylindrica* gametocidal chromosome. *Genes Genet Syst* 72: 243–248
- Shizuya H, Birren B, Kim U-J, et al. (1992) Cloning and stable maintenance of 300-kilobase fragments of human DNA in *Escherichia coli* using and f-factor-based vector. *Proc Natl Acad Sci USA* 89: 8794–8797
- Soderlund C, Humphray S, Dunham, A et al. (2000) Contigs built with fingerprints, markers, and FPC v4.7. *Genome Res* 10: 1772–1787
- Soderlund C, Longden I, Mott R (1997) FPC: A system for building contigs from restriction fingerprinted clones. *Comp Appl Biosci* 13: 523–535
- Stein N, Graner A (2004) Map-based gene isolation in cereal genomes. In: Gupta P, Varshney R (eds.), *Cereal genomics*. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp 331–360

- Stephens J L, Brown S E, Lapitan N L V, et al. (2004) Physical mapping of barley genes using an ultrasensitive fluorescence in situ hybridization technique. *Genome* 47: 179–189
- Sulston J, Mallett F, Durbin R, et al. (1989) Image analysis of restriction enzyme fingerprint autoradiograms. *Comp Appl Biosci* 5: 101–106
- Sulston J, Mallett F, Staden R, et al. (1988) Software for genome mapping by fingerprinting techniques. *Comp Appl Biosci* 4: 125–132
- Telenius H, Carter N P, Bebb C E, et al. (1992) Degenerate oligonucleotide-primed pcr: General amplification of target DNA by a single degenerate primer. *Genomics* 13: 718–725
- Thangavelu M, James A B, Bankier A, et al. (2003) HAPPY mapping in a plant genome: Reconstruction and analysis of a high-resolution physical map of a 1.9 Mbp region of *Arabidopsis thaliana* chromosome 4. *Plant Biotech J* 1: 23–31
- The International Human Genome Mapping C (2001) A physical map of the human genome. *Nature* 409: 934–941
- Valarik M, Bartos J, Kovarova P, et al. (2004) High-resolution FISH on super-stretched flow-sorted plant chromosomes. *Plant J* 37: 940–950
- Varshney R K, Grosse I, Haehnel U et al. (2006) Genetic mapping and bac assignment of EST-derived SSR markers shows non-uniform distribution of genes in the barley genome. *Theor Appl Genet* 113: 239–250
- Wang C-J R, Harper L, Cande W Z (2006) High-resolution single-copy gene fluorescence in situ hybridization and its use in the construction of a cytogenetic map of maize chromosome 9. *Plant Cell* 18: 529–544
- Wardrop J, Fuller J, Powell W, et al. (2004) Exploiting plant somatic radiation hybrids for physical mapping of expressed sequence tags. *Theor Appl Genet* 108: 343–348
- Wardrop J, Snape J, Powell W, et al. (2002) Constructing plant radiation hybrid panels. *Plant J* 31: 223–228
- Waugh R, Dear P H, Powell W et al. (2002) Physical education – new technologies for mapping plant genomes. *Trends Plant Sci* 7: 521–523
- Wei F, Coe E, Nelson W, et al. (2007) Physical and genetic structure of the maize genome reflects its complex evolutionary history. *PLoS Genet* 3: e123
- Wendl M C, Waterston R H (2002) Generalized gap model for bacterial artificial chromosome clone fingerprint mapping and shotgun sequencing. *Genome Res* 12: 1943–1949
- Wong G K S, Yu J, Thayer E C et al. (1997) Multiple-complete-digest restriction fragment mapping: Generating sequence-ready maps for large-scale DNA sequencing. *Proc Natl Acad Sci USA* 94: 5225–5230
- Wu C, Sun S, Lee M-K, et al. (2005) Whole-genome physical mapping: An overview on methods for DNA fingerprinting. In: Meksem K, Kahl G (eds.), *The handbook of plant genome mapping. Genetic and physical mapping*. WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim, pp 257–306
- You F M, Luo M-C, Gu Y Q, et al. (2007) Genoprofiler: Batch processing of high-throughput capillary fingerprinting data. *Bioinformatics* 23: 240–242
- Yu Y, Tomkins J P, Waugh R, et al. (2000) A bacterial artificial chromosome library for barley (*Hordeum vulgare* L.) and the identification of clones containing putative resistance genes. *Theor Appl Genet* 101: 1093–1099
- Zhang L, Cui X, Schmitt K, et al. (1992) Whole genome amplification from a single cell: Implications for genetic analysis. *Proc Natl Acad Sci* 89: 5847–5851
- Zhou S, Bechner M C, Place M, et al. (2007) Validation of rice genome sequence by optical mapping. *BMC Genomics* 8: 278

Chapter 12

Map-Based Cloning of Genes in Triticeae (Wheat and Barley)

Simon Krattinger, Thomas Wicker, and Beat Keller

Abstract In the Triticeae crops wheat and barley, only genetic information is available for many genes underlying agronomically important traits, as well as for natural or induced mutants. Therefore, map-based (positional) cloning is the most promising approach for the molecular isolation of genes causing these traits or mutant phenotypes. A growing number of genes (currently 19) have been isolated from wheat and barley based on genetic information only, and a lot has been learnt on the most suitable approaches for such challenging projects in the large and complex genomes of Triticeae. With the ongoing or starting projects on the generation of high-resolution physical maps in barley and wheat, map-based cloning will become simpler and faster in the near future. In order to fully exploit these new resources, there is an increasing need for high-resolution mapping populations. In addition, new and efficient tools have to be developed for the validation of the candidate genes identified in positional cloning. Here, we review the state of the art of positional gene cloning in the Triticeae crops and discuss the challenges in this field of research in the next years.

12.1 Introduction

Wheat and barley genetically behave as diploid plants. Wheat is allohexaploid with the three homoeologous genomes A, B and D ($2n = 42$), each derived from a diploid ancestor species, whereas barley is a classical diploid species ($2n = 14$). However, wheat chromosomes from the three different genomes do not pair randomly during meiosis but specifically with their homologous chromosomes, resulting in a genetically diploid behavior of genes in segregation studies. Some genes like *Ph1* prevent pairing of homoeologous chromosomes in wheat (see also Chapter 8). Both wheat and barley have a long and successful history of genetic analysis both in basic research as well as plant breeding. This has

B. Keller (✉)

Institute of Plant Biology, University of Zürich, Zürich, Switzerland
e-mail: bkeller@botinst.unizh.ch

resulted in long lists of genetically mapped traits (e.g. see McIntosh et al. 2007 for wheat; <http://wheat.pw.usda.gov/ggpages/bgn/26/text261a.html#9> for barley). To understand these traits at the molecular level, it is essential to isolate the underlying genes and to study their mode of action in the plant. Positional cloning in wheat and barley has remained a challenge mainly because of (i) problems related to the enormous genome sizes and (ii) the underdeveloped genomic resources for efficient cloning. The genome sizes of 5×10^9 bp for barley and 1.6×10^{10} bp for wheat make them by far the largest genomes where map-based cloning has been successfully used in the absence of extensive physical maps or complete genome sequences. The genomes contain high contents of repetitive DNA, mostly retrotransposons, but it seems that an unexpectedly large part of the genome also consists of DNA transposons (Wicker et al. 2003). The repetitive nature of these genomes makes steps like chromosome walking and marker development inherently difficult. Interestingly, although it was thought that the hexaploid nature of the wheat genome would make positional cloning even more difficult because of the difficulty to distinguish the three homoeologs, this has not turned out to be a too substantial problem. In fact, the three genomes are very different in the intergenic regions and it has been relatively simple to “stay” on the correct chromosome during chromosome walking. In addition, the fact that BAC libraries of diploid or tetraploid model genomes have become available has further reduced the problem of hexaploidy for positional cloning. Despite the problems of genome size and repetitive genome content, a growing number of genes has been isolated from wheat and barley using positional cloning. Mostly, disease resistance genes have been cloned, but there are more and more genes for other traits which have been isolated. Among others, this includes the cloning of the first QTL in wheat (Uauy et al. 2006).

12.2 Genes Isolated from Wheat and Barley by Positional Cloning

Map-based cloning is a strategy to clone genes of interest without prior knowledge of the gene product. In the following, we want to give an overview on the challenges for map-based cloning in Triticeae as well as the successful strategies to circumvent them.

Figure 12.1 provides an overview of the procedure to clone a gene of interest via map-based cloning. A prerequisite for map-based cloning is a suitable fine-mapping population derived from a cross between two parents differing in the trait of interest. Combination of phenotypic scoring and molecular marker data lead to a precise genetic map indicating the position of the gene of interest. To get the corresponding physical information, the two closest flanking markers are used to screen BAC libraries. Chromosome walking as well as sequencing of the target interval will lead to the identification of candidate genes.

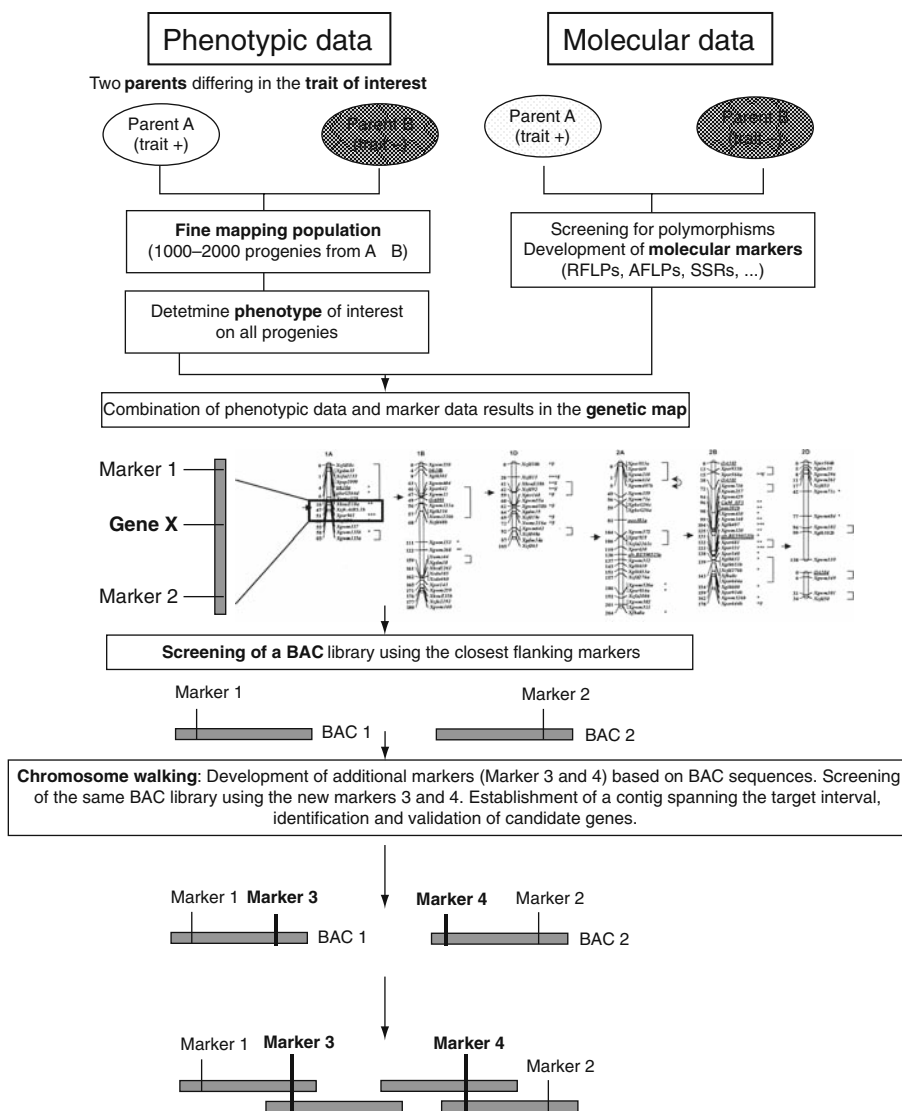


Fig. 12.1 Strategy to isolate a gene of interest (gene X) via map-based cloning

A summary of the genes cloned to date and the corresponding references are given in Table 12.1. Although the number of genes is still limited, there is great progress both in the type of genes that have been cloned since the first gene was isolated by map-based cloning in barley (*mlo*, Buschges et al. 1997) as well as in the efficiency of the approaches.

Nearly half of the cloned genes in Triticeae are involved in disease resistance against biotrophic pathogens. Many of them encode proteins with conserved

Table 12.1 List of genes that have been positionally cloned in wheat and barley

Gene name	Gene function	Population size	Size of genetic and physical interval		Candidate gene validation	References
			[cM/kb] and Mb/cM ratio			
<i>Vrs1</i> (barley)	HD-ZIP transcription factor, responsible for two-rowed spikes in barley	9,831 gametes	0.07 cM/518 kb (7.4 Mb/cM)		expression studies, mutant analysis, allele diversity study	Komatsuda et al. (2007)
<i>Bot1</i> (barley)	666-amino acid protein with 10–12 transmembrane helices conferring boron-toxicity tolerance	6,720 gametes	0.15 cM (md)		homology to Arabidopsis, expression studies, growth tests in yeast	Sutton et al. (2007)
<i>Ppd-H1</i> (barley)	Pseudo response regulator involved in control of flowering	4,672 gametes	0.04 cM (md)		allele diversity study, mutant analysis, expression studies	Turner et al. (2005)
<i>Gpc-B1</i> (wheat)	NAC transcription factor controlling senescence and grain protein, zinc, and iron content	9,000 gametes	7.4 kb (md)		allele diversity study, RNA interference, expression studies	Uauy et al. (2006)
<i>Q</i> (wheat)	AP2 transcription factor influencing threshing character and spike phenotype	930 gametes of chromosome substitution lines	0.7 cM/~250 kb (0.36 Mb/cM)		homology to Arabidopsis, mutation analysis, allele diversity study, expression analysis	Faris et al. (2003)
<i>VRN1</i> (wheat)	AP1 like MADS-box transcription factor controlling flowering	6,190 gametes	0.04 cM/~550 kb (13.5 Mb/cM)		homology to Arabidopsis, allelic variation, expression studies	Yan et al. (2003)
<i>VRN2</i> (wheat)	Dominant repressor of flowering that is down regulated by vernalization	5,698 gametes	0.04 cM/439 kb (11 Mb/cM)		expression studies, allele diversity study, RNA interference	Yan et al. (2004)

Table 12.1 (continued)

Size of genetic and physical interval [cM/kb] and Mb/cM ratio				
Gene name	Gene function	Population size	Candidate gene validation	References
<i>VNR3</i> (wheat and barley)	VNR3 is an orthologue of Arabidopsis Flowering Time (FT)	high density genetic map in barley based on 1,600 gametes	allele diversity study, expression profile, transformation	Yan et al. (2006)
<i>Ph1</i> (wheat)	Multigene family, heterochromatin or both. A major chromosome pairing locus in polyploid wheat	2.5 Mb (md)	deletion mutants	Griffiths et al. (2006)
Resistance Genes				
<i>Mla</i> (barley)	CC-NBS-LRR Powdery mildew resistance	3,600 gametes	mutant analysis, genetic complementation in a single cell transient expression assay, transgenic plants	Wei et al. (1999), Zhou et al. (2001), Halterman et al. (2001), Halterman et al. (2003), Shen et al. (2003), Halterman and Wise (2004)
<i>Pm3</i> (wheat)	CC-NBS-LRR Powdery mildew resistance	2,680 gametes	mutation analysis, expression studies, complementation in a single cell transient expression assay	Yahiaoui et al. (2004), Srichumpa et al. (2005), Yahiaoui et al. (2006)
<i>mlo</i> (barley)	Transmembrane protein Powdery mildew resistance	4,044 gametes	mutant analysis	Buschges et al. (1997)
<i>Lr1</i> (wheat)	CC-NBS-LRR Leaf rust resistance	~7,300 gametes	transgenic plants, virus induced gene silencing	Qiu et al. (2007), Cloutier et al. (2007)

Table 12.1 (continued)

Gene name	Gene function	Population size	Size of genetic and physical interval		Candidate gene validation	References
			[cM/kb] and Mb/cM ratio			
<i>Lr10</i> (wheat)	CC-NBS-LRR Leaf rust resistance	6,240 gametes	0.13 cM/300 kb (2.3 Mb/cM)		mutant analysis, allele diversity study, transgenic plants	Stein et al. (2000), Wicker et al. (2001), Feuillet et al. (2003)
<i>Lr21</i> (wheat)	NBS-LRR Leaf rust resistance	1,040 gametes	1.7 cM (md)		transgenic plants	Huang et al. (2003)
<i>Rpg1</i> (barley)	Tandem kinase Stem rust resistance	8,518 gametes	0.02 cM/110 kb (5.5 Mb/cM)		allele diversity study, transgenic plants	Brueggeman et al. (2002), Horvath et al. (2003)
<i>rym4/rym5</i> (barley)	Eukaryotic translation factor 4E Resistance against <i>Barley Yellow Mosaic</i> and <i>Barley Mild Mosaic Virus</i>	7,772 gametes	650 kb (md)		allele diversity study, transgenic plants	Stein et al. (2005)
<i>Ror2</i> (barley)	Syntaxin required for basal penetration resistance against powdery mildew	1,522 gametes	0.26 cM (md)		mutant analysis	Collins et al. (2003)
<i>Rar1</i> (barley)	25.5 kDa protein, convergence point in the signaling of resistance to powdery mildew	~8,000 gametes	300 kb (md)		mutant analysis	Shirasu et al. (1999a)

md: Missing data. A Mb/cM ratio can not be provided

domains such as nucleotide binding site (NBS) and leucine rich repeat (LRR) domains (*Mla*, *Pm3*, *Lr1*, *Lr10*, *Lr21*). With the exception of *Lr21*, these genes confer race-specific resistance. In barley, there is a second and highly unusual recessive resistance against powdery mildew based on the absence of a functional *Mlo* gene. The result is a broad spectrum resistance which is active against most powdery mildew races in barley and has been durable for a long time. *Rym4/rym5* is an additional recessive resistance gene that has been identified in the barley gene pool and that confers resistance against Barley Yellow Mosaic and Barley Mild Mosaic Virus.

Besides resistance genes, some of the genes identified by map-based cloning contribute to important and complex agronomic traits such as grain protein content, vernalization requirement, photoperiod sensitivity, tolerance to abiotic stress and morphological and domestication traits. These genes often encode transcription factors and they are probably involved in the regulation of the expression of downstream genes.

12.3 Genetic Mapping

The availability of a high resolution genetic map is the basis for the successful isolation of a desired gene. This is a first and very time-consuming step. Fine-mapping requires an accurate population that has to be saturated with molecular markers. In a first step, a low-resolution map containing hundreds of markers all over the genome is established using only a small number of plants (100–200). There are thousands of molecular markers available for the wheat and barley genomes (e.g. Roder et al. 1998; Paillard et al. 2003), allowing an efficient and rapid localisation of genes or mutations of interest at low resolution (see also Chapter 9). The resolution of such populations is around 1–5 cM. To get a suitable high-resolution genetic map, thousands of plants are screened with the closest flanking markers identified on the low-resolution population. Only plants that are recombinant between the flanking markers are selected for further saturation with molecular markers and for phenotyping. The gene responsible for six-rowed spikes in barley for example was mapped to a 0.07 cM target interval screening a fine-mapping population consisting of more than 4,900 plants (Komatsuda et al. 2007).

A marker defines a specific chromosomal fragment or locus within the genome. According to Falconer and Mackay (1996), the ideal type of genetic marker should be (i) abundant, (ii) neutral with respect to the trait of interest and to reproductive fitness, (iii) co-dominant, and (iv) highly polymorphic. Molecular markers are based on differences on the DNA and emerged in the 1980s. RFLPs (Restriction Fragment Length Polymorphisms) were the first molecular markers to be used. The polymorphism detected by this marker type involves recognition of specific restriction sites and hybridization with specific probes. Complete genetic maps have been constructed using RFLP markers

(example in barley: Graner et al. 1991). With the emergence of the Polymerase Chain Reaction (PCR) – technology, PCR based molecular markers such as RAPDs (Random Amplified Polymorphic DNAs), STS (Sequence Tagged Sites), AFLPs (Amplified Fragment Length Polymorphisms), and SSRs (Simple Sequence Repeats) provided new, cheaper, and faster tools to investigate the genome. The availability of large EST databases as well as technological progress allowed the establishment of Single Nucleotide Polymorphism markers (SNPs) (<http://probes.pw.usda.gov:8080/snpworld/Search>) (See Chapter 9).

Allopolyploid organisms with large genome sizes like wheat represent an exceptional challenge to develop molecular markers. The lack of genome sequence information complicates development of PCR and SNP based markers. Furthermore, PCR markers derived from genes tend to amplify sequences from homoeologous chromosomes and RFLP probes usually hybridize to more than one genome. Hence molecular markers often map to undesired positions in homoeologous genomes. To solve this problem and to facilitate genetic mapping in wheat, highly useful cytogenetic stocks have been developed. The polyploid nature of modern bread wheat allowed the construction of more than 400 deletion lines derived from the cultivar “Chinese Spring” (Endo and Gill 1996). Each line has a specific deletion which is detectable by chromosome analysis. Furthermore, a collection of chromosome substitution lines, where a particular chromosome is substituted by a chromosome of a different cultivar, has been constructed. Mapping populations derived from chromosome substitution lines ideally differ only in one particular chromosome carrying the gene of interest. All polymorphic markers will therefore map to the desired chromosome and mapping of polymorphic markers to homoeologous regions can be excluded. Faris and colleagues used a combination of deletion and substitution lines to map the major domestication gene *Q*. The parents of their mapping population differed only in chromosome 5A that contained the gene of interest. With a population size of only 465 F2 plants they were able to clone this gene of great agronomic importance (Faris et al. 2003).

Many agronomically important traits are quantitative and genetically complex, with several or many genes (Quantitative Trait Loci QTL) involved in the specific phenotypes. However, there is occasionally a single QTL among the genes responsible for a trait which contributes a large part of 30% or more to the phenotypic variance of the trait of interest, allowing a specific targeting of the underlying gene by map-based cloning. The isolation of such major QTLs requires special mapping populations. The most frequent approach is the development of Near Isogenic Lines (NILs), where the gene of interest from one parent is crossed into the genetic background of the other parent of the population. This allows to genetically remove by segregation all other loci contributing to the trait which are genetically not closely linked.

12.4 Physical Mapping for Map-Based Cloning

Successful isolation of a gene requires the establishment of a physical region of DNA spanning the target interval that was identified by genetic mapping. The closest molecular markers identified by genetic mapping are used to screen large insert libraries. In the early time of positional cloning Yeast Artificial Chromosome (YAC) libraries were used. They were soon replaced by Bacterial Artificial Chromosome (BAC) libraries with insert sizes of ~ 80 to >150 kb. Gaps in physical maps have to be closed by chromosome walking or by PCR. Since the ratios of genetic to physical distances are highly variable within a genome, it is difficult to estimate the physical distance and the number of BAC clones between two flanking markers. Yan et al. (2004) reported a ratio between genetic and physical distance of ~ 1.7 Mb/cM during isolation of the flowering repressor VRN2 in wheat. It is thought that gene dense regions have a higher recombination frequency than long stretches of repetitive elements. At the leaf rust resistance locus *Lr10*, the recombination frequencies changed from 0.6 Mb/cM to 12 Mb/cM within a 230 kb physical distance (Stein et al. 2000). Usually, a target interval of <0.5 cM is required to be able to establish a suitable physical contig.

When the first gene was cloned in Triticeae (*mlo*, Buschges et al. 1997) there were no YAC or BAC libraries available yet. Generation of a physical contig was very laborious at that time since a large YAC library with more than 40,000 clones had to be developed specifically for that project. Nowadays many BAC libraries of barley, hexaploid wheat as well as several wild grasses are available. Some of them, for example the BAC library of *Ae. tauschii* spp. *strangulata*, accession AL8/78 (Luo et al. 2003), have been fingerprinted and arranged in contigs using automated high-throughput procedures thereby providing physical maps and accelerating map based cloning.

Currently, many new resources are created that will be highly valuable for map-based cloning efforts in the future. In the frameworks of the International Wheat Genome Sequencing Consortium (IWGSC) and the International Barley Genome Sequencing Consortium (IBSC), high-quality genetic maps are integrated with BAC contigs, partially generated from chromosome-specific libraries (see Chapters 10 and 11). Although the ultimate goal of complete physical maps for barley and wheat may lie several years in the future, the genomic resources are of use for map-based cloning already now. A table listing all available wheat, barley and rye BAC libraries is provided in Chapter 9.

12.5 Application and Problems of Chromosome Walking in Triticeae

The conceptually straight-forward approach of cloning a target gene by chromosome walking is burdened with a number of problems and difficulties when used to clone genes in Triticeae: Because of the large size of the Triticeae

genomes, BAC libraries are very large. To obtain a 6-fold coverage of the barley or diploid wheat genome (e.g. *Triticum monococcum*), approximately 275,000 BAC clones with an average insert size of 120 kb are needed. BAC libraries of tetraploid or hexaploid wheat species are accordingly larger. The screening of such a library with radioactive RFLP probes (the classical method) is very labour-intensive. More convenient is the screening by PCR, if multi-dimensional pools of BAC clones are available.

A major problem for physical mapping is the uneven coverage of genomes by BAC libraries. The usual 6-fold coverage of the genome provides an approximately 98% chance that all regions are covered, assuming an even distribution of BAC clones. Therefore, gaps are found in every BAC library. If a chromosome walking effort reaches a gap in the library, it comes to a halt. The only way to securely bridge such a gap is to use a different (or expanded) BAC library. During the cloning of the powdery mildew gene *Pm3* in wheat, for example, chromosome walking in the diploid wheat *T. monococcum* was impossible because of a gap in the BAC library. Chromosome walking and isolation of the *Pm3* gene was only successful after a different BAC library from the tetraploid wheat *T. turgidum* had been used (Yahiaoui et al. 2004).

An additional complication of map-based cloning in polyploid wheat is the presence of three closely related, homoeologous subgenomes, A, B, and D. Special care must be given to identify specifically the BAC clones of the genome of interest. It is recommended to test each important BAC for its chromosomal location by deriving single copy markers from the BAC which can then be located in cytogenetic stocks (e.g. deletion lines). Once genome-specific markers have been developed, and BAC clones from the correct subgenome have been isolated, subgenome chromosome walking is usually relatively unproblematic, because the three subgenomes differ strongly in their repeat architecture.

12.6 Problems Caused by Repetitive Elements

Chromosome walking is usually performed using probes derived from BAC end sequences. However, more than 80% of the Triticeae genomes consist of repetitive DNA (see Chapters 14 and 15) and it can be very problematic to obtain specific probes from a given BAC end. For example, if the end of a BAC clone contains a highly repetitive element of the BARE1-group, it is virtually impossible to use this sequence for screening of a BAC library with a radioactive probe derived from that end. It still might be possible to produce a specific PCR probe, but success is by no means guaranteed. In a few cases it was possible to derive a specific low- or single-copy probe from a Triticeae BAC end (e.g. Qiu et al. 2007). However, this must be considered the exception rather than the rule. The recent development of an approach to derive specific markers spanning the junction of two transposable elements is greatly supporting the identification of new PCR markers for chromosomal walking (Paux et al. 2006).

If no single-copy probe can be generated from a BAC end, a repetitive one has to be used for a chromosome walking step. Such an approach can be extremely labour-intensive. During the cloning of the *Lr10* gene in wheat, one repetitive probe had to be used which identified more than 100 BAC clones in the library (Stein et al. 2000). The correct clones could only be identified by isolating DNA from all identified BACs and performing fingerprint analysis. The fingerprint pattern proved to be specific enough for the identification of BAC clones that overlapped with the ones from which the probe was generated and the chromosome walking step could be completed.

12.7 Aspects of Sequencing and Identification of Candidate Genes

Once a BAC contig that contains flanking genetic markers (and thus, the entire target region) is established, the search for candidate genes begins. If the contig consists of several BAC clones and/or resources are limited, an initial round of low-pass shotgun sequencing is advised. In this approach, a shotgun library from all BAC clones of the contig is constructed (as would be done for complete sequencing), but only a small number of clones is sequenced. Usually, sequencing of 96 shotgun clones in forward and reverse direction provides sufficient sequence information that the entire gene (and repeat) content of a 120 kb BAC clone can be reliably sampled. The low-pass sequence reads are then used for BLAST searches against various databases such as rice protein, Triticeae repeats or simply entire GenBank. If this analysis yields only hits to known repeat sequences (i.e. transposable elements), the BAC does not have to be sequenced further as it most likely contains no genes.

If genic sequences are detected, the BAC clone is usually chosen for complete sequencing, in order to obtain the complete sequences of all genes, pseudogenes and gene fragments present on the BAC. Once a complete sequence is available, all potential gene sequences (i.e. everything that was not characterised as a known repeat sequence) have to be considered candidates for the gene of interest, as they all lie in the interval between the flanking genetic markers. At this stage, it has also to be considered from what type of variety the BAC library used for physical analysis was derived from. In the ideal case, it was made from a line containing the gene/QTL of interest and we can be sure that the gene to be cloned is present on the BAC. However, frequently such a BAC library will not be available and one always has to consider the possibility that the gene of interest is not present on the analysed BAC sequences. In this case, which is particularly frequent for race-specific resistance genes, other strategies such as selective PCR amplification of regions of interest or specific cosmid or fosmid libraries from the line containing the gene have to be considered.

As described in Chapter 15, complete sequencing of a BAC clone can be a challenging and labour-intensive task. Sequencing the entire target interval (low-pass or complete) can be costly, because it sometimes can span several

BACs from one flanking marker to the other. Target intervals in Triticeae tend to be so large because recombination occurs mainly in gene-containing regions. Thus, BAC contigs often contain extended regions comprised exclusively of transposable elements. Despite their large physical size, the genetic distance between markers in highly repetitive regions might be minimal or even zero. During the cloning of the *HvEF4E* gene, a region of more than 300 kb consisting exclusively of transposable elements was included in the BAC contig and completely sequenced (Wicker et al. 2005).

12.8 The Use and Limits of Model Genomes for Marker Development and Map-Based Cloning in Triticeae

As described above, the development of molecular markers is complicated by the large genome size of Triticeae and the absence of genome sequence information. Despite a great variation in genome size, all grasses show an extensive conservation of gene order along chromosome segments (Gale and Devos 1998). This so-called synteny reflects their descent from a common ancestor grass approximately 50 million years ago (Paterson et al. 2004). Synteny can be exploited for marker development, since the order of genetic markers (if these are conserved coding genes) along the chromosome is also largely conserved. Model organisms like rice or *Brachypodium* that have a small genome and show good synteny to wheat have been widely used to generate molecular markers in wheat and barley. The rice genome is completely sequenced and is therefore the ideal anchor genome for map-based cloning in other grasses. In a typical approach, genetic markers that are close to the gene of interest in Triticeae can be used to identify the corresponding region in the rice genome. Genes from that region in rice can then be used to generate new markers for mapping to narrow down the interval in which the target gene is located. They can also be used to screen a BAC library and isolate the region of interest from a Triticeae genome. Thus, the availability of a model genome can greatly facilitate the targeted development of new genetic markers in other grasses. *ROR2*, which is required for basal penetration resistance in barley against powdery mildew, was successfully isolated using a barley-rice syntenic map-based cloning approach (Collins et al. 2003).

The main problem for the use of a model genome is caused by the fact that gene collinearity between grasses has many exceptions. Previous studies have shown that the number collinear genes can vary greatly from <50 to >80% in different genomic regions (Guyot et al. 2004; Lagudah et al. 2006; Bossolini et al. 2007; Faris et al. 2008), but often the genes of interest are not in syntenic positions. The best examples are resistance genes of the NBS-LRR type. These are of great economic importance but are notorious for not being found in syntenic regions in grasses. In fact, mechanisms have been postulated that indicate a selective advantage if defence-related genes are moved around in

the genome frequently (Leister 2004). Colinearity is therefore useful for the identification and isolation of a target region but candidate genes have to be carefully defined and worked with on a case-by-case basis.

In recent years, *Brachypodium distachyon* has emerged as a new model genome for Triticeae. The *Brachypodium* genome is relatively small and a first sequence assembly indicating a genome size of approximately 300 Mbp is available (see Chapter 16 and www.brachypodium.org). It was also shown that *Brachypodium* is clearly more closely related to Triticeae than is rice. It was estimated that the divergence of the *Brachypodium* lineage from the Triticeae has occurred approximately 35 million years ago (Bossolini et al. 2007). This estimate was obtained under the assumption that the rice/Triticeae divergence occurred about 50 million years ago. The use of *Brachypodium* as a model genome was successful for the analysis of a region containing the major chromosome pairing locus in polyploid wheat, *Ph1*. This region was identified using synteny information from *Brachypodium sylvaticum* (Griffiths et al. 2006).

Despite the closer relationship of *Brachypodium* and Triticeae, numerous differences in colinearity were detected (Fig. 12.2, Bossolini et al. 2007). The overall order of genetic markers between *Brachypodium* and wheat was

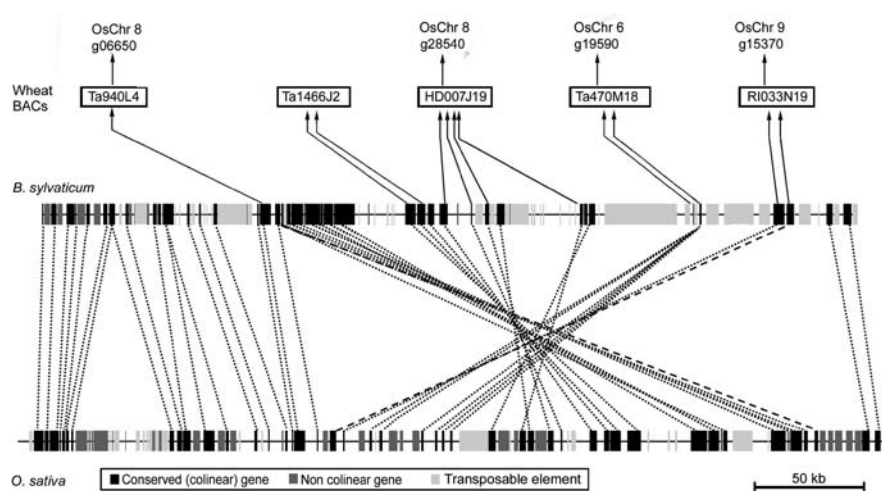


Fig. 12.2 Comparison of the *Lr34* locus from wheat (*top*) *Brachypodium sylvaticum* (*middle*) and rice (*bottom*). Genes conserved between *Brachypodium* and rice are depicted as black boxes whereas non-conserved genes are depicted as dark grey boxes. Orthologous genes are connected by dotted lines. Transposable elements are indicated as light grey boxes. From the *Lr34* region from wheat only five BAC clones (indicated as boxes with BAC addresses) were available which cover parts of the locus. Solid arrows indicate homologs of *Brachypodium* genes found on wheat BACs. On four wheat BACs, genes were found whose putative rice orthologs (labelled with TIGR gene identifiers, very top) map to different rice chromosomes. Adapted from Bossolini et al. (2007)

perfectly conserved but showed a large inversion compared to rice. In addition, low-pass sequencing of wheat BAC clones from the target region showed that there are many differences in gene content between *Brachypodium* and wheat. Thus, *Brachypodium* was clearly shown to be more similar to Triticeae than rice. Nevertheless, the numerous differences observed demonstrate that, despite the availability of the *Brachypodium* genome sequence, complete sequencing of large regions in Triticeae will be essential.

12.9 Validation of Candidate Genes

High resolution genetic and physical mapping can delineate a gene or a QTL to a small region of the genome. In most cases, more than one candidate gene will be identified. Therefore, the final step of map-based cloning consists of the careful analysis of all candidate genes present in the physical target interval to identify the gene or the regulatory region underlying the trait. This task is frequently underestimated in terms of the required effort and depends on the size of the physical target interval and the gene density. In the last few years, new tools have become available for gene validation in wheat and barley. Expression data, allele diversity studies, mutant analysis, virus-induced gene silencing and stable transformation are experimental approaches to demonstrate the correlation between specific sequences in the target interval and the investigated trait.

Gene density in Triticeae is very variable. Gene dense islands are separated from each other by repeat-rich regions containing only few coding sequences. Two extremes are represented by (i) a 518 kb physical target interval spanning the *Vrs1* locus responsible for six-rowed barley with only one functional gene (Komatsuda et al. 2007) and by the *VRN2* target interval with a gene density of one coding sequence per 55 kb (Yan et al. 2004). Thus, the number of candidate genes can vary dramatically, even if the physical interval to be studied might be very similar in size. Some traits, for example race-specific disease resistance, allow a good prediction of the gene to be expected. Such resistance is frequently based on genes encoding NBS-LRR type of proteins. The identification of a single NBS-LRR coding gene in the genetic interval carrying the resistance gene makes it likely that this is actually the gene of interest. Nevertheless, such a simple situation is rather exceptional, and in many cases more sophisticated approaches are needed.

In the ideal case, functional null mutants are available. Sequence analysis of candidate genes in these mutants can rapidly reveal whether the candidate gene is the target gene. The barley *mlo* powdery mildew resistance gene was quickly confirmed this way, as well as the barley *Mla1* and *Mla12* alleles (Shen et al. 2003; Zhou et al. 2001), the *Lr10* leaf rust resistance gene in wheat (Feuillet et al. 2003) and the wheat powdery mildew resistance gene *Pm3b* (Yahiaoui et al. 2004).

In the case of the powdery mildew pathogen of wheat and barley, the host pathogen interaction occurs exclusively in epidermal cells which can be transiently transformed by particle bombardment (Douchkov et al. 2005; Schweizer et al. 1999; Shirasu et al. 1999b). Using marker genes such as genes encoding beta-glucuronidase or green fluorescent protein, the transformed cells can be identified. By cobombardment with a candidate resistance gene and subsequent infection with a specific race of the powdery mildew pathogen, transformed cells can be examined for compatible or incompatible interaction with the pathogen.

This approach allows to test whether a candidate gene confers resistance and has been successfully used for the *mlo* gene (Shirasu et al. 1999b) and the *Mla* alleles (Halterman et al. 2001; Shen et al. 2003; Zhou et al. 2001) in barley as well as for the *Pm3* allelic series in wheat (Srichumpa et al. 2005; Yahiaoui et al. 2006; Yahiaoui et al. 2004). Such transient assays are of course extremely valuable but are currently restricted to young seedlings and the epidermal tissue.

Stable transformation in cereals such as wheat and barley as well as maize is now routine but still tedious and time consuming before results can be obtained in the T1 or T2 generation. In barley, stable *Agrobacterium*-mediated transformation is mostly restricted to the cultivar Golden Promise, but it was also successfully used in the line "Igri" (Stein et al. 2005). In contrast, particle bombardment is still the most widely used method in wheat. There, several sublines of cultivar Bobhwhite (Feuillet et al. 2003; Simons et al. 2006), the Canadian cultivar Fielder (Cloutier et al. 2007) as well as a number of other cultivars (including the winter wheat cultivar Jagger, Yan et al. 2006) were stably transformed in different projects. Thus, there is still a restricted availability of lines which can be efficiently transformed, but the list of suitable lines has been expanding in the last years.

Recently, virus-induced gene silencing (VIGS) based on the Barley Stripe Mosaic Virus (BSMV) has been developed as a new tool for functional gene analysis in barley and wheat (Holzberg et al. 2002; Scofield et al. 2005). In this system, short fragments (up to 1,200 nucleotides) are cloned into the γ -genome of the tripartite viral genome. After infection of the plants at the young seedling stage (barley or wheat), the target gene is effectively silenced. In the case of a candidate gene actually being the resistance gene, we expect silencing of the resistance gene (and susceptibility) after infection of seedling genotypes carrying the resistance gene of interest with the relevant pathogen race. The power of VIGS in cereals was shown for the wheat *Lr21* leaf rust resistance gene which was efficiently silenced by VIGS and seedlings became susceptible (Scofield et al. 2005). This confirmed the results previously obtained by stable transformation (Huang et al. 2003). Very recently, the *Lr1* leaf rust resistance gene could also be confirmed this way (Cloutier et al. 2007). Thus, candidate genes whose function can be determined at the seedling stage can be efficiently silenced and particularly resistance genes active at the seedling stage can be identified. There is a possible problem with this system which has to be considered in its

application: in the barley gene pool there is widespread resistance against BSMV, preventing its use in these lines. Therefore, for each barley line to be used for VIGS, it has to be established first that it is susceptible to the virus. In wheat cultivars, resistance to BSMV is rare or absent, making VIGS more straightforward (S. Bieri and B. Keller, unpublished data).

It was recently shown that RNAi based gene silencing is effective against related, homoeologous genes in wheat (Travella et al. 2006). RNAi was successfully used for gene validation of the *VRN2* gene in wheat (Yan et al. 2004). Although RNAi is a very efficient strategy, it has to be kept in mind that it involves the time-consuming generation of stably transformed plants, with the limits of genotypes that can be transformed.

Finally, haplotype analysis, gene expression data as well as the study of allelic diversity can give strong hints for the identification of the correct candidate gene. Although a definitive proof is hard to achieve this way, such studies can greatly support map-based cloning. Haplotype studies were successfully used for the cloning of *VRN1* (Yan et al. 2003), *VRN2* (Yan et al. 2004), *VRN3* (Yan et al. 2006) and *Ppd-H1* (Turner et al. 2005), to give just a few examples.

12.10 The Role of Bioinformatics in Map-Based Cloning

Bioinformatics plays a crucial role in several steps of map-based cloning. An important aspect is the design of molecular markers from model genomes. As described above, the region collinear to the target Triticeae region is isolated from a model genome (e.g. rice) and genes present there are considered for possible COS (conserved orthologous sequences) marker development (see Chapter 17). The bioinformatics tasks range from excision of the collinear region from the model genome to automated download and annotation of all coding sequences in that region to the identification of Triticeae EST sequences that correspond to those genes. It is crucial in this approach that genes annotated in the collinear region of the model genome are examined very carefully.

During chromosome walking, BAC clones have to be sampled for their gene and repeat content based on data from low-pass shotgun sequencing. This requires hundreds of sequence reads to be searched for the presence of known repeats and potential gene sequences. Additionally, the quality of the shotgun library has to be assessed by determining the degree of contamination with *E. coli* genomic DNA. These goals are achieved by running BLAST searches against various databases. It would be very time-consuming to run and view all the hundreds of BLAST searches manually. Thus, programs are necessary that automatically run the searches and screen the outputs for specific hits to databases. For example, the content of *E. coli* DNA can be estimated by simply screening all BLASTN outputs for strong hits with genomic *E. coli* sequences (e.g. alignments of hundreds of bp with very strong DNA identity, or very low

E-values). This approach allows a rapid assessment of the quality, the repeat and gene content of a particular shotgun library.

A BAC clone which is chosen for complete sequencing is covered with shotgun sequence reads to a minimum of a 6-fold coverage, usually even 8–10-fold. The raw shotgun reads are assembled into sequence contigs with broadly available software such as PhredPrap (Ewing et al. 1998), leading to a phase-1 unfinished sequence which consists of multiple (or even dozens) of unordered pieces. The unordered sequence contigs have to be brought into their correct linear order, to allow closing gaps by specific PCR. This second phase of BAC sequencing is very time-consuming and labor-intensive. TE sequences often contain problematic motifs that cause gaps in the sequences. Additionally, high-copy repeats (such as BARE1 or Angela retrotransposons) can cause mis-assemblies if several copies of such elements are present on a BAC clone. As most sequence gaps are in one way or another caused by TEs, the latter have to be studied intensely in the phase 2 of BAC sequencing. Target site duplications which flank almost all TEs and nesting patterns provide important information to infer the linear order of sequence contigs based on which primers can then be designed to bridge the gaps.

In summary, it is evident that TEs are the cause of problems at all levels of map-based cloning in Triticeae. They are an obstacle in designing specific genetic markers for mapping and chromosome walking and they are the major cause for problems during sequencing. Thus, a solid knowledge of the most abundant TE families and a high-quality database of repetitive elements are essential for improvements of the efficiency of map-based cloning in Triticeae.

12.11 Outlook

The improvement of genomic resources (mainly physical maps) for application in Triticeae promises to make map-based cloning of genetically tractable genes, mutations and QTL a topic of increasing importance in the near future. The emphasis in map-based cloning in the Triticeae in the next years should be on the efficient use of the genomic resources that are and will be developed. Enormous genomic resources are currently being created by the International Wheat and Barley Genome Sequencing Consortia. They will facilitate map-based cloning dramatically. Although association mapping will certainly become a highly relevant tool for gene identification in wheat, populations specifically targeting genes of interest will remain essential and should be vigorously developed, considering the time-frame of such work. After many years of genetic analysis, we are now in the position to advance to an understanding of the molecular basis of many of the specific traits found in the Triticeae crops. This is not only a very exciting time for science, but also a challenge to apply the new knowledge for practical applications and improvements in breeding strategies.

References

- Bossolini, E., Wicker, T., Knobel, P.A. and Keller, B. (2007) Comparison of orthologous loci from small grass genomes Brachypodium and rice: implications for wheat genomics and grass genome annotation. *Plant J.* 49, 704–717.
- Brueggeman, R., Rostoks, N., Kudrna, D., Kilian, A., Han, F., Chen, J., Druka, A., Steffenson, B. and Kleinbols, A. (2002) The barley stem rust-resistance gene *Rpg1* is a novel disease-resistance gene with homology to receptor kinases. *Proc. Natl. Acad. Sci. USA* 99, 9328–9333.
- Buschges, R., Hollricher, K., Panstruga, R., Simons, G., Wolter, M., Frijters, A., van Daelen, R., van der Lee, T., Diergaarde, P., Groenendijk, J., Topsch, S., Vos, P., Salamini, F. and Schulze-Lefert, P. (1997) The barley *mlo* gene: a novel control element of plant pathogen resistance. *Cell* 88, 695–705.
- Cloutier, S., McCallum, B.D., Loutre, C., Banks, T.W., Wicker, T., Feuillet, C., Keller, B. and Jordan, M.C. (2007) Leaf rust resistance gene *Lr1*, isolated from bread wheat (*Triticum aestivum* L.) is a member of the large psr567 gene family. *Plant Mol. Biol.* 65, 93–106.
- Collins, N.C., Thordal-Christensen, H., Lipka, V., Bau, S., Kombrink, E., Qiu, J.L., Huckelhoven, R., Stein, M., Freialdenhoven, A., Somerville, S.C. and Schulze-Lefert, P. (2003) SNARE-protein-mediated disease resistance at the plant cell wall. *Nature* 425, 973–977.
- Douchkov, D., Nowara, D., Zierold, U. and Schweizer, P. (2005) A high-throughput gene-silencing system for the functional assessment of defense-related genes in barley epidermal cells. *Mol. Plant Microbe Int.* 8, 755–761.
- Endo, T.R. and Gill, B.S. (1996) The deletion stocks of common wheat. *J. Hered.* 87, 295–307.
- Ewing, B., Hillier, L., Wendl, M.C. and Green, P. (1998) Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res.* 8, 175–185.
- Falconer, D.S. and Mackay, T.F.C. (1996) *Introduction to Quantitative Genetics*. Longman, Essex, England.
- Faris, J.D., Fellers, J.P., Brooks, S.A. and Gill, B.S. (2003) A bacterial artificial chromosome contig spanning the major domestication locus *Q* in wheat and identification of a candidate gene. *Genetics* 164, 311–321.
- Faris, J.D., Zhang, Z., Fellers, J.P. and Gill, B.S. (2008) Micro-colinearity between rice, Brachypodium, and *Triticum monococcum* at the wheat domestication locus *Q*. *Funct. Integr. Genomics* [Epub ahead of print].
- Feuillet, C., Travella, S., Stein, N., Albar, L., Nublat, A. and Keller, B. (2003) Map-based isolation of the leaf rust disease resistance gene *Lr10* from the hexaploid wheat (*Triticum aestivum* L.) genome. *Proc. Natl. Acad. Sci. USA* 100, 15253–15258.
- Gale, M.D. and Devos, K.M. (1998) Plant comparative genetics after 10 years. *Science* 282, 656–659.
- Graner, A., Jahoor, A., Schondelmaier, J., Siedler, H., Pillen, K., Fischbeck, G., Wenzel, G. and Herrmann, R.G. (1991) Construction of an RFLP map of barley. *Theor. Appl. Genet.* 83, 250–256.
- Griffiths, S., Sharp, R., Foote, T.N., Bertin, I., Wanous, M., Reader, S., Colas, I. and Moore, G. (2006) Molecular characterization of *Ph1* as a major chromosome pairing locus in polyploid wheat. *Nature* 439, 749–752.
- Guyot, R., Yahiaoui, N., Feuillet, C. and Keller, B. (2004) In silico comparative analysis reveals a mosaic conservation of genes within a novel colinear region in wheat chromosome 1AS and rice chromosome 5S. *Funct. Integr. Genomics* 4, 47–58.
- Halterman, D.A., Zhou, F.S., Wei, F.S., Wise, R.P. and Schulze-Lefert, P. (2001) The Mla6 coiled-coil, NBS-LRR protein confers AvrMla6-dependent resistance specific to *Blumeria graminis* f.sp. *hordei* in barley and wheat. *Plant J.* 25, 335–348.
- Halterman, D.A., Wei, F.S. and Wise, R.P. (2003) Powdery mildew-induced *Mla* mRNAs are alternatively spliced and contain multiple upstream open reading frames. *Plant Physiol.* 131, 558–567.

- Halterman, D.A. and Wise, R.P. (2004) A single-amino acid substitution in the sixth leucine-rich repeat of barley *Mla6* and *Mla13* alleviates dependence on *Rar1* for disease resistance signaling. *Plant J.* 38, 215–226.
- Holzberg, S., Brosio, P., Gross, C. and Pogue, G.P. (2002) Barley stripe mosaic virus-induced gene silencing in a monocot plant. *Plant J.* 3, 315–327.
- Horvath, H., Rostoks, N., Brueggeman, R., Steffenson, B., von Wettstein, D. and Kleinhofs, A. (2003) Genetically engineered stem rust resistance in barley using the *Rpg1* gene. *Proc. Natl. Acad. Sci. USA* 100, 364–369.
- Huang, L., Brooks, S.A., Li, W., Fellers, J.P., Trick, H.N. and Gill, B.S. (2003) Map-based cloning of leaf rust resistance gene *Lr21* from the large and polyploid genome of bread wheat. *Genetics* 164, 655–664.
- Komatsuda, T., Pourkheirandish, M., He, C., Azhaguvel, P., Kanamori, H., Perovic, D., Stein, N., Graner, A., Wicker, T., Tagiri, A., Lundqvist, U., Fujimura, T., Matsuoka, M., Matsumoto, T. and Yano, M. (2007) Six-rowed barley originated from a mutation in a homeodomain-leucin zipper I-class homeobox gene. *Proc. Natl. Acad. Sci. USA* 104, 1424–1429.
- Lagudah, E.S., McFadden, H., Singh, R.P., Huerta-Espino, J., Bariana, H.S. and Spielmeier, W. (2006) Molecular genetic characterization of the *Lr34/Yr18* slow rusting resistance gene region in wheat. *Theor. Appl. Genet.* 114, 21–30.
- Leister, D. (2004) Tandem and segmental gene duplication and recombination in the evolution of plant disease resistance genes. *Trends Genet.* 20, 116–122.
- Luo, M.C., Thomas, C.S., Deal, K.R., You, F.M., Anderson, O.D., Gu, Y.Q., Li, W., Kuraparthi, V., Gill, B., McGuire, P.E. and Dvorak, J. (2003) Construction of contigs of *Ae. tauschii* genomic DNA fragments cloned in BAC and BiBAC vectors. In: N.E. Ponga, M. Romano, E.A. Ponga and G. Galterio (Eds.), *Proceedings of the 10th International Wheat Genetics Symposium*. Istituto Sperimentale per la Cerealicoltura, Rome.
- McIntosh, R.A., Devos, K.M., Dubcovsky, J., Rogers, W.J., Morris, C.F., Appels, R., Somers, D.J. and Anderson, O.A. (2007) Catalogue of gene symbols for wheat: 2007 supplement. In: W.R. Raupp (Ed.), *Annual Wheat Newsletter*, Volume 53. Kansas State University, Manhattan, KS, pp. 159–180.
- Paillard, S., Schnurbusch, T., Winzeler, M., Messmer, M., Sourdille, P., Aberhalden, O., Keller, B. and Schachermayr, G. (2003) An integrative genetic linkage map of winter wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* 107, 1235–1242.
- Paterson, A.H., Bowers, J.E. and Chapman, B.A. (2004) Ancient polyploidization predating divergence of the cereals, and its consequences for comparative genomics. *Proc. Natl. Acad. Sci. USA* 101, 9903–9908.
- Paux, E., Roger, D., Badaeva, E., Gay, G., Bernard, M., Sourdille, P. and Feuillet, C. (2006) Characterizing the composition and evolution of homoeologous genomes in hexaploid wheat through BAC-end sequencing on chromosome 3B. *Plant J.* 48, 463–474.
- Qiu, J.W., Schurch, A.C., Yahiaoui, N., Dong, L.L., Fan, H.J., Zhang, Z.J., Keller, B. and Ling, H.Q. (2007) Physical mapping and identification of a candidate for the leaf rust resistance gene *Lr1* of wheat. *Theor. Appl. Genet.* 115, 159–168.
- Roder, M.S., Korzun, V., Wendehake, K., Plaschke, J., Tixier, M.H., Leroy, P. and Ganal, M.W. (1998) A microsatellite map of wheat. *Genetics* 149, 2007–2023.
- Schweizer, P., Pokorny, J., Aberhalden, O. and Dudler, R. (1999) A transient assay system for the functional assessment of defense-related genes in wheat. *Mol. Plant Microbe Int.* 12, 647–654.
- Scofield, S.R., Huang, L., Brandt, A.S. and Gill, B.S. (2005) Development of a virus-induced gene-silencing system for hexaploid wheat and its use in functional analysis of the *Lr21*-mediated leaf rust resistance pathway. *Plant Physiol.* 138, 2165–2173.
- Shen, Q.H., Zhou, F., Bieri, S., Haizel, T., Shirasu, K. and Shulze-Lefert, P. (2003) Recognition specificity and RAR1/SGT1 dependence in barley *Mla* disease resistance genes to the powdery mildew fungus. *Plant Cell* 15, 732–744.

- Shirasu, K., Lahaye, T., Tan, M.W., Zhou, F., Azevedo, C. and Schulze-Lefert, P. (1999a) A novel class of eukaryotic zinc-binding proteins is required for disease resistance signaling in barley and development in *C. elegans*. *Cell* 99, 355–366.
- Shirasu, K., Nielsen, K., Piffanelli, P., Oliver, R. and Schulze-Lefert, P. (1999b) Cell-autonomous complementation of mlo resistance using a biolistic transient expression system. *Plant J.* 17, 293–299.
- Simons, K.J., Fellers, J.P., Trick, H.N., Zhang, Z.C., Tai, Y.S., Gill, B.S. and Farris, J.D. (2006) Molecular characterization of the major wheat domestication gene *Q*. *Genetics* 172, 547–555.
- Srichumpa, P., Brunner, S., Keller, B. and Yahiaoui, N. (2005) Allelic series of four powdery mildew resistance genes at the *Pm3* locus in hexaploid bread wheat. *Plant Physiol.* 139, 885–895.
- Stein, N., Feuillet, C., Wicker, T., Schlagenhauf, E. and Keller, B. (2000) Subgenome chromosome walking in wheat: a 450-kb physical contig in *Triticum monococcum* L. spans the *Lr10* resistance locus in hexaploid wheat (*Triticum aestivum* L.). *Proc. Natl. Acad. Sci. USA* 97, 13436–13441.
- Stein, N., Perovic, D., Kumlehn, J., Pello, B., Stracke, S., Streng, S., Ordon, F. and Graner, A. (2005) The eukaryotic translation initiation factor 4E confers multiallelic recessive Bymovirus resistance in *Hordeum vulgare* (L.). *Plant J.* 42, 912–922.
- Sutton, T., Baumann, U., Hayes, J., Collins, N.C., Shi, B.J., Schnurbusch, T., Hay, A., Mayo, G., Pallotta, M., Tester, M. and Langridge, P. (2007) Boron-toxicity tolerance in barley arising from efflux transporter amplification. *Science* 318, 1446–1449.
- Travella, S., Klimm, T.E. and Keller, B. (2006) RNA interference-based gene silencing as an efficient tool for functional genomics in hexaploid bread wheat. *Plant Physiol.* 142, 6–20.
- Turner, A., Beales, J., Faure, S., Dunford, R.P. and Laurie, D.A. (2005) The pseudo-response regulator Ppd-H1 provides adaptation to photoperiod in barley. *Science* 310, 1031–1034.
- Uauy, C., Distelfeld, A., Fahima, T., Blechl, A. and Dubcovsky, J. (2006) A NAC gene regulating senescence improves grain protein, zinc, and iron content in wheat. *Science* 314, 1298–1301.
- Wei, F.S., Gobelman-Werner, K., Morroll, S.M., Kurth, J., Mao, L., Wing, R., Leister, D., Schulze-Lefert, P. and Wise, R.P. (1999) The *Mla* (powdery mildew) resistance cluster is associated with three NBS-LRR gene families and suppressed recombination within a 240-kb DNA interval on chromosome 5S (1HS) of barley. *Genetics* 153, 1929–1948.
- Wicker, T., Stein, N., Albar, L., Feuillet, C., Schlagenhauf, E. and Keller, B. (2001) Analysis of a contiguous 211 kb sequence in diploid wheat (*Triticum monococcum* L.) reveals multiple mechanisms of genome evolution. *Plant J.* 26, 307–316.
- Wicker, T., Guyot, R., Yahiaoui, N. and Keller, B. (2003) CACTA transposons in Triticeae – a diverse family of high-copy repetitive elements. *Plant Physiol.* 132, 52–63.
- Wicker, T., Zimmermann, W., Perovic, D., Paterson, A.H., Ganal, M., Graner, A. and Stein, N. (2005) A detailed look at 7 million years of genome evolution in a 439 kb contiguous sequence at the barley *Hv-eIF4E* locus: recombination, rearrangements and repeats. *Plant J.* 41, 184–194.
- Yahiaoui, N., Srichumpa, P., Dudler, R. and Keller, B. (2004) Genome analysis at different ploidy levels allows cloning of the powdery mildew resistance gene *Pm3* from hexaploid wheat. *Plant J.* 37, 528–538.
- Yahiaoui, N., Brunner, S. and Keller, B. (2006) Rapid generation of new powdery mildew resistance genes after wheat domestication. *Plant J.* 47, 85–98.
- Yan, L., Loukoianov, A., Tranquilli, G., Helguera, M., Fahima, T. and Dubcovsky, J. (2003) Positional cloning of the wheat vernalization gene *Vrn1*. *Proc. Natl. Acad. Sci. USA* 100, 6263–6268.
- Yan, L., Loukoianov, A., Blechl, A., Tranquilli, G., Ramakrishna, W., SanMiguel, P., Bennettzen, J.L., Echenique, V. and Dubcovsky, J. (2004) The wheat *Vrn2* gene is a flowering repressor down-regulated by vernalization. *Science* 303, 1640–1644.

- Yan, L., Fu, D., Li, C., Blechl, A., Tranquilli, G., Bonafede, M., Sanchez, A., Valarik, M., Yasuda, S. and Dubcovsky, J. (2006) The wheat and barley vernalization gene *Vrn3* is an orthologue of FT. *Proc. Natl. Acad. Sci. USA* 103, 19581–19586.
- Zhou, F.S., Kurth, J.C., Wei, F.S., Elliott, C., Vale, G., Yahiaoui, N., Keller, B., Somerville, S., Wise, R. and Schulze-Lefert, P. (2001) Cell-autonomous expression of barley *Mla1* confers race-specific resistance to the powdery mildew fungus via a Rar1-independent signaling pathway. *Plant Cell* 13, 337–350.

Chapter 13

Functional Validation in the *Triticeae*

Ingo Hein, Jochen Kumlehn, and Robbie Waugh

Abstract The need to discover and confirm gene function directly in both wheat and barley is growing in importance. In this chapter we outline three of the most common approaches – TILLING, ‘transient’ and ‘stable’ transformation – that are being adopted by *Triticeae* researchers to meet this objective. As these approaches have different outcomes, they are therefore applicable to different situations depending upon the research questions being addressed. Here, we summarise recent developments in each of these functional validation strategies, and where possible use examples to illustrate the power of each approach.

13.1 Introduction

The rapid and successful development of genomics resources, technologies and information focused on *Triticeae* species has resulted in a situation where a bottleneck in gene discovery is frequently the functional validation of candidate genes. As the key to function generally lies in a species-specific biological context, it is crucial that functional validation strategies are developed within the framework of the crop under investigation. Here, we outline a number of current experimental systems that address this requirement. We have focused our discussion on approaches that have recently been adopted or developed by the *Triticeae* research community to link genes to functions and genotypes to phenotypes. These systems exploit a combination of natural and induced variation along with transient and stable genetic transformation. We have provided examples from our own experience or from the literature that highlight the power of each approach in a given experimental situation. In all cases, the starting point is a gene or DNA sequence that on the balance of experimental evidence already collected is involved in a specific biological process. The DNA sequence can come from our own research activity or from the

I. Hein (✉)

Scottish Crop Research Institute, Invergowrie, Dundee, DD2 5DA, UK
e-mail: ingo.hein@scri.ac.uk

literature and the key objective of the following approaches is to investigate the link between this piece of DNA sequence and a specific biological function.

13.2 Targeted Induced Local Lesions in Genomes (TILLING)

In 1928, L.J. Stadler reported the successful induction of mutations in crop plants (barley and maize) using x-rays and radium and outlined the potential of mutations as both a fundamental and applied tool in genetic research (Stadler 1928a,b, originals not consulted). Seventy-two years later the reverse genetics procedure known as TILLING was described (McCallum et al. 2000a,b). TILLING couples the use of saturating chemical (or potentially physical) mutagenesis, with sensitive post-PCR sequence-specific mutation detection procedures. TILLING employs pooled DNAs assembled from large populations of plants allowing the efficient identification of individuals containing mutations in target amplicons. Since its development, TILLING has rapidly become a general and effective platform for 'reverse genetics' in a wide range of plants (Till et al. 2004; Perry et al. 2004; Dalmay et al. 2008) including the *Triticeae* cereals. Several populations have been developed for barley (Caldwell et al. 2004; Talamè et al. 2008; Nils Stein (NS) pers. comm.; Peter Langridge (PL) pers. comm.) and wheat (Slade et al. 2005; Christobal Uauy (CU) and Jorge Dubcovsky (JD) pers. comm.; Peter Sharp (PS) pers. comm.; Andy Phillips (AP) pers. comm.). However, at the time of writing, the majority remain unpublished. In these populations, mutations that represent heritable changes in target genes are being identified and incorporated into gene-function validation strategies and in applied programmes where they provide biological resources that can be rapidly adopted by plant breeders for crop improvement. A key to their widespread adoption is that once TILLING populations have been developed they require no 'special' containment facilities and face no regulatory hurdles to their widespread growth or cultivation.

13.2.1 Mutagens and Mutation Frequency

TILLING populations are generated by mutagenesis. Mutagens are forms of energy or chemical substances that increase the frequency of mutations above that of background in the genomes of exposed organisms. Since the 1920s, several mutagens have been used to induce genetic variation and to develop mutant populations. For TILLING, ethyl methane sulphonate (EMS) has become the most commonly used because of its ease of handling, well established mode of action and its effectiveness in inducing a high frequency of point mutations in the absence of gross chromosomal abnormalities (Ashburner 1990). EMS adds alkyl groups to the hydrogen-bonding oxygen of guanine producing O-6 alkylguanine which pairs with T (instead of C) and causes G/C

to A/T transitions. These are the type of mutations almost exclusively recovered from TILLING populations (Greene et al. 2003; Dalmais et al. 2008). However, most recently Talamè et al. (2008) described the use of NaAzide as a mutagen for barley TILLING population development and *N*-methyl-*N*-nitrosourea (MNU) was used for rice (Suzuki et al. 2008).

Because it is of paramount importance to the quality of the resource, optimising the mutation frequency in a TILLING population is essential. This must be determined empirically: if too low, too many plants will be required to discover mutations in a target gene and if too high, viability and/or sterility will likely be a problem. Mutation frequency has commonly been estimated on the basis of phenotypic screens for characters like embryonic lethality (seed set), seedling lethality, chlorophyll deficiency or single copy gene phenotypes. This approach does not however provide an accurate measure for many reasons that include functional redundancy and the general complexity of biological systems. At the DNA sequence level there are no direct measures available to globally assess mutation frequency, perhaps outside of the 'next generation sequencing' methodologies. Amplified Fragment Length Polymorphism (AFLP) has been used to estimate mutation frequency induced by tissue culture in somaclonal variants (Matthes et al. 2001) and to estimate mutation frequency in the Barley *Dis*-TILLING population described by Caldwell et al. (2004). However, as only homozygous loss and both homozygous and heterozygous addition of bands can be scored, AFLP scanning is at best qualitative providing only an estimate of mutation frequency.

Performing a mutation screen and confirming the nature of induced alleles by re-sequencing directly determines the effective mutation frequency for a target amplicon. Once this has been done several times, mutation frequencies have been found to range from 1 per 24 Kb reported in the original publication describing TILLING in hexaploid wheat (Slade et al. 2005) to less than 1 per 500 kb described in barley (Caldwell et al. 2004). Similar frequencies have been described subsequently in other wheat (CU and JD, AP, PS, pers. comm) and barley (Talamè et al. 2008; NS pers. comm.) populations. At these frequencies TILLING populations can be a manageable size and still contain sufficient variation at any given locus to be useful. In wheat, 246 induced mutations were identified in the waxy locus (granule bound starch synthase I) in a population of only 1,920 plants. The experience of several labs that have subsequently developed TILLING populations in hexaploid wheat is similar. Tolerance to such a high mutation frequency is almost certainly attributable to the buffering capacity of the polyploid wheat genome against deleterious mutations. The relative ease with which mutations can be discovered in polyploid wheat is however balanced by two things. First, the amount of crossing and selection required to remove background mutations and second, functional redundancy (i.e. many single genes have functional orthologues on all three chromosomes). In such cases it may be necessary to combine homozygous disruptions in all homoeologous alleles before being able to test for phenotype.

In barley, because of the lower frequency of mutations, working populations are necessarily larger, generally in the order of 4000–10,000 plants (Caldwell et al. 2004; Talamè et al. 2008; NS, PL pers. comm.). The number of induced alleles recovered from a screen is conversely lower than that achieved in wheat and, being diploid, the association of mutation with phenotype relatively straightforward.

13.2.2 Mutation Spectrum Analysis

Mutations in DNA can be categorized as silent, mis-sense or truncation events according to how the encoded protein is affected. In the species where sufficient information is available on mutation spectrum, the observed distribution of the different types of event has been very close to the expected distribution. In *Arabidopsis*, truncation events were observed 3.6 times as frequently in heterozygotes suggesting that the homozygote has significantly impaired fitness (Greene et al. 2003). A similar situation is expected in the *Triticeae*, with the exception being the polyploids where buffering may allow a higher frequency of gene knockouts. Furthermore, a local compositional bias in the nucleotides flanking individual mutations was also observed in *Arabidopsis* but there was no evidence for EMS-induced mutation hotspots. Published reports and personal communications generally support and extend these observations to the data currently emerging from the *Triticeae*.

13.2.3 Web-Based Computational Tools for TILLING

Computational tools that can assist in amplicon choice for TILLING, for primer design and for assessment of the likely impact of discovered mutations on protein function have been developed and made accessible via proWeb (<http://www.proweb.org/#proWeb>). CODDLE (Codons Optimised to Discover Deleterious LEsions) (<http://www.proweb.org/coddle/>) helps identify regions of genes that are likely to be important for protein function and once a defined region has been selected suitable primers can be designed using Primer3 (<http://frodo.wi.mit.edu/>). We have found that primer design and amplification condition optimization are crucial, and there is a strong correlation between amplicon quality and TILLING success. Gelbuddy software (<http://www.proweb.org/gelbuddy/>) helps rapidly extract information from TILLING gels and speeds up gel analysis. After induced alleles have been identified and characterized, two further web based utilities, SIFT (Sorting Intolerant from Tolerant) <http://blocks.fhcrc.org/sift/SIFT.html> (Ng and Henikoff 2003) and PARSESNP (Project Aligned Related Sequences and Evaluate SNPs) <http://www.proweb.org/parsesnp/> (Taylor and Greene 2003) help predict the effect that a particular amino acid substitution might have on

protein function (amongst other things). While all of these tools have not been universally adopted, they are potentially useful filters for prioritizing mutant alleles for further study.

13.2.4 Populations for Reverse Genetics

For TILLING, high quality structured populations are needed in order to track any discovered mutations back to the families from which they originated and to avoid or control for complications due to contamination through outcrossing. Importantly, the initial seed stock should be genetically pure. As the material that will be mutagenized is usually a multi cellular structure like a seed, and mutational events are single cell specific, the generation of plants grown from mutagenized seed (known as the M1) will be chimeric (i.e. will contain sectors of cells containing a different spectrum of mutations). The chimeric M1 plants contain two classes of mutation: somatic and genetic. Only the genetic mutations are heritable as they are present in the germ line of mutagenized individuals. The number of cells in the germ line (the *initials*) that will ultimately contribute to the next generation of seed has been termed the Genetically Effective Cell Number (GECN, Li and Redei 1969). Mutated *initials* are the basis of the population that forms the fundamental biological TILLING resource. As the cereals may already contain several shoots in the embryo of the mature grain, and thus multiple sets of initials, the GECN is higher than that of 2 estimated for *Arabidopsis*. In barley, Hodgdon et al. (1981) estimated the GECN as six. This means that five sixths of the progeny from an individual M1 plant would likely be derived from a cell in which a specific gene (*A*) is not mutated (i.e. is *AA*), and one sixth from a cell that has been mutated (*Aa*). The segregation ratio for the double recessive *aa* genotype in the M2 progeny from such an M1 plant would therefore be 23:1 (i.e. $(5 \times AA) + AA + Aa + aA:aa$). The somatic mutations are eradicated by purification through seed and are not observed in the M2 plants (though they may be observed in seed on the M1 plants which contains maternal tissue). To minimise the occurrence of the same mutation being represented multiple times in a TILLING population, a single seed from each individual M1 plant is therefore taken forward to the M2 generation. This practice ensures that each plant in the population originates from a uniquely different mutagenized cell. Figure 13.1 outlines the approach adopted in both wheat and barley to generate TILLING populations.

The template for mutation detection is DNA isolated from each of the individual M2 and the essential biological resource is the seed collected from exactly the same plants. Bagging the individual M2 to prevent cross contamination should be considered seriously, if practicable. Phenotyping the entire population is then possible using seed harvested from the M2 plants (the M3). In diploid barley, a considerable amount of phenotypic variation will be

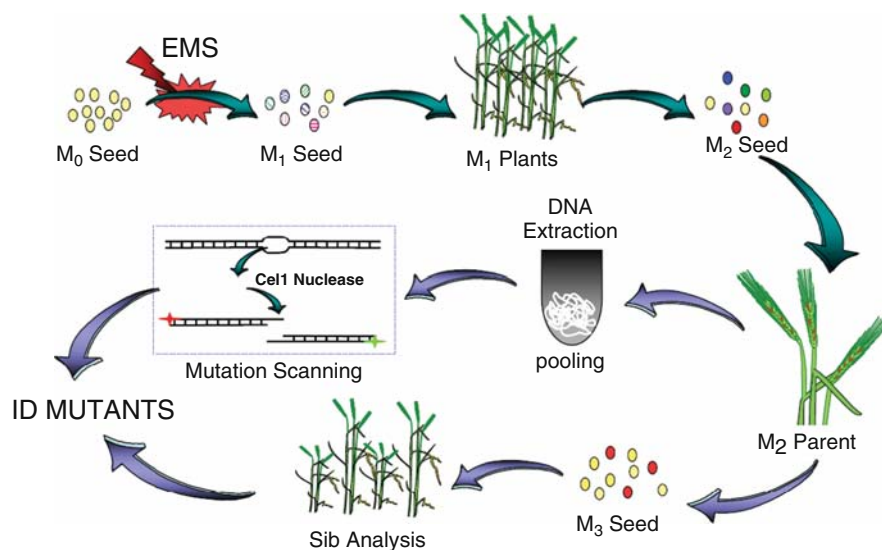


Fig. 13.1 Development of a TILLING resource for the *Triticeae*. Normal seed (M₀) are mutagenized to create the M₁ and these seed (M₁ seed) are planted to generate chimeric M₁ plants. A single M₂ seed taken from each M₁ plant will contain an unique spectrum of mutations. These will be both homozygous and heterozygous. DNA extracted from individual M₂ plants is the resource for mutation scanning. Seed from the same plants (M₃ seed) is the biological resource for validation and further exploitation. Much of the phenotypic variation is released in M₃ families as recessive alleles, heterozygous in the M₂, will segregate in the M₃. TILLING for mutations employs pooled DNA from several individual M₂ plants. Associating mutations with phenotypes is then achieved by co-segregation of mutant alleles with an associated phenotype in the M₃ families (See Color Insert)

released in the M₃ as recessive mutations that were heterozygous in the M₂ plants segregate in a 3:1 ratio. In tetraploid and hexaploid wheat the release of phenotypic variation is largely masked because of complementation by homeo-loci on the homeologous chromosomes (AP, CU & JD, PS pers. comm.)

13.2.5 Mutation Detection and Validation

For mutation detection, M₂ DNAs are generally pooled to a depth of between 4 and 8 individuals (8–16 alleles). Target specific PCRs on each of the pooled DNAs will amplify all alleles in the population. If a mutation is present in the sequence of one of the component amplicons of a pool, denaturing the mixed PCR product and allowing the products to re-anneal will form heteroduplexes. The presence of heteroduplexes (i.e. mutations) can be identified by a number of approaches (summarised in Waugh et al. 2006). Currently the most common methodology employs a DNA mismatch cleavage assay using an endonuclease purified from celery called CEL I. CEL I cleaves heteroduplex dsDNA at single

base mismatches generating two cleavage products. CEL I is distributed by Transgenomic under the trademark SurveyorTM nuclease (Qui et al. 2004) but effective protocols for its purification have been published (Yang et al. 2000) and use of 'home made' enzyme is commonplace.

PCR amplification of target sequences for TILLING generally employs two different fluorescently-labelled primers. The precise labels are dictated by the fragment detection platform used. In the presence of a mutant allele CEL I cleavage generates shorter fragments that are separated from the uncleaved product by gel or capillary electrophoresis. The LI-COR fluorescence-based fragment analysis system combined with IR700 and IR800 labelled amplicons (or DY-681 and DY-781 from Biomers.net) is probably the most common approach for mutation detection. The system is rapid, high throughput and relatively tolerant of PCR amplicon quality. As CEL I cleavage generates different sized and coloured fragments that together total the size of the uncleaved amplicon, potential false positives that do not add up to the size of the initial amplicon are easily excluded and the physical location of the mutation can be inferred. In addition to LI-CORs, fragment analysis is also being conducted using capillary-based DNA sequencers (Jonathan Clark, JIC pers. comm.) and straightforward agarose gel electrophoresis of unlabelled amplicons and ethidium bromide staining (CU and JD, PS pers. comm.).

13.2.6 Mutation Confirmation and Functional Validation

Once a DNA pool containing a putative mutation has been identified, re-sequencing alleles from the individuals used to make the pools will confirm the presence, location and nature of an induced mutation and its consistency with the mode of action of the mutagen used. The link between induced mutation and phenotype can then be assessed by sibling analysis of heterozygous M3 families directly or in selfed families of homozygous M3 plants backcrossed to the original parent. Extensive backcrossing or further population development is seldom required to prove the link between the target mutation and function and the need to remove background mutations is generally not required – at least to assess the potential phenotypic impact of a gene specific mutation. However, the high mutational load in polyploid wheat is a potential concern when attempting to correlate mutant alleles with phenotypic effects. In this case, the detection of multiple independent mutational events is of significant value in the validation process.

The plant research community is currently expressing considerable interest in using induced mutations as a bridge between fundamental knowledge gained in models like *Arabidopsis* and rice into applications in a broader suite of crops. As it is now realistic to discover stable and heritable genetic lesions in effectively any gene of interest directly in a crop; linking genes to functions is achievable in a species-specific biological context. Furthermore, it is certain that some of the

induced variation will have significant commercial or agronomic potential. For example, modifying the properties of cereal starches (Slade et al. 2005) and generating *de novo* resistance to important agricultural pathogens (Büschges et al. 1997; Kanyuka et al. 2005) emphasises how this rapid, effective and uncontroversial route to the generation of biological diversity promises much for both basic and applied research.

13.3 Transient Gene Validation Assays

13.3.1 Virus Induced Gene Silencing (VIGS)

Virus induced gene silencing (VIGS) is a powerful reverse genetics tool used in plants to functionally validate host candidate genes. The underlying mechanism is based on post-transcriptional silencing of host genes by homologous DNA fragments that are transiently expressed via a suitable virus system. Activation of the host's sequence-specific RNA degradation mechanism not only targets the RNAs of the viral genome for degradation, but also the transcribed plant sequences introduced by the recombinant virus (Ruiz et al. 1998; Moissiard and Voinnet 2006). Numerous viral systems have been described for VIGS studies in dicotyledonous plants. However, to date, only two such approaches have been reported for monocotyledonous plants and are based on *Barley stripe mosaic virus* (BSMV) and *Brome mosaic virus* (BMV) systems (Holzberg et al. 2002; Ding et al. 2006).

BSMV: BSMV, a member of the *Hordeivirus* family, is a tripartite, positive-sense RNA virus that infects several plant species including barley, oats, wheat and maize (McKinney and Greeley 1965). The three single-stranded genomic RNAs present in this virus, referred to as α , β and γ , encode seven major proteins and one translational read-through protein (Johnson et al. 2003). BSMV strains including ND18, from which BSMV VIGS vectors are derived, have been shown to be seed transmissible, a *Hordeivirus* characteristic which has been associated with the function of the protein encoded by γ b (Edwards 1995). Indeed, seed transmissibility and maintenance of silencing in progeny of silenced plants has been shown recently (Bruun-Rasmussen et al. 2007). Furthermore, it has been demonstrated that other BSMV seed-transmitted strains, such as MI-1, invade reproductive meristems very early in their development (Johansen et al. 1994).

Fusions of the green fluorescent protein (GFP) to the C-terminal end of γ b protein result in systemic expression of GFP and provide a phenotypic marker to follow systemic virus spread (Haupt et al. 2001). Phloem unloading of GFP was first observed from major longitudinal veins, although very patchy along their length, and finally into mesophyll and epidermis cells. Furthermore, these results indicate that the C-terminal end of γ b protein is suitable to insert host target genes for VIGS applications. Host-specific barley and wheat sequences in

the form of sense, antisense or hairpin constructs have been successfully introduced downstream of the γ b protein, which has been modified by adding a stop codon to prevent translation of the inserts *in planta* or interference with γ b protein activity (Holzberg et al. 2002; Scofield et al. 2005; Lacomme et al. 2003).

In barley, BSMV vectors with a deletion of the coat protein ($\Delta\beta$ a) have enhanced ability to silence plant endogenous genes such as phytoene desaturase (*Pds*), a gene required for synthesizing carotenoids, which are required to prevent chlorophyll from photo-bleaching (Brigneti et al. 2004; Holzberg et al. 2002). However, this modification can result in greater necrosis accompanying the photobleaching observed in *Pds* silenced barley plants and the enhanced silencing response of the $\Delta\beta$ a-modified virus remains elusive in wheat (Bruun-Rasmussen et al. 2007). Other studies by Hein et al. (2005) and Bruun-Rasmussen et al. (2007) have shown that the silencing response and BSMV-induced symptoms can vary greatly in a cultivar-specific manner, highlighting the need to assess silencing efficacy and BSMV infection symptoms prior to carrying out VIGS experiments for functionally characterizing candidate genes.

The photobleaching phenotype of *Pds* silenced barley plants can typically be observed as early as 7 days post inoculation (dpi) in systemically-infected leaves of the main tiller, albeit significant silencing of endogenous *Pds* mRNA occurs up to two days earlier (Bruun-Rasmussen et al. 2007). However, newly formed leaves after 21 dpi on the main tiller typically show wild-type BSMV symptoms and no longer demonstrate photobleaching (Holzberg et al. 2002). Interestingly, Holzberg et al. (2002) and Hein et al. (2005) have shown that virus symptoms and silencing occurs in second tillers until at least 21 dpi. The transient nature of the silencing response is thought to be due to partial or complete loss of host inserts via recombination occurring during virus replication and is associated with insert stability (Bruun-Rasmussen et al. 2007). In plants that express a *Pds* silencing response through the entire life span and, furthermore, in second generation plants, the *Pds* insert, which was originally 400 bp in size was reduced to 80–125 nt in size, following virus recombination (Bruun-Rasmussen et al. 2007).

To date, many VIGS studies using BSMV as a silencing vector are aimed at unravelling the function of candidate genes in disease resistance. For example, the requirement of *Rar1*, *Sgt1* and *Hsp90* in *Mla13*-based resistance in barley towards the powdery mildew pathogen *Blumeria graminis* f.sp. *hordei* (*Bgh*) could be demonstrated (Hein et al. 2005), as could their role and the function of *Lr21* in *Lr21*-mediated resistance towards leaf rust, caused by *Puccinia triticina* in hexaploid wheat (Scofield et al. 2005). Nevertheless, BSMV-based functional validation of candidate genes is not limited to disease resistance assays. Silencing of *P23k*, which has a cellular distribution coinciding with cell wall polysaccharides, resulted in abnormal leaf development, asymmetric orientation of main veins, and cracked leaf edges caused by mechanical weakness. These observations led to a proposed function of *P23k* in the synthesis of cell wall polysaccharides and contribution to secondary wall formation in barley leaves (Oikawa et al. 2007).

BMV: The second virus that has been described more recently as a suitable VIGS vector in barley, rice and maize is BMV, a member of the *Bromovirus* genus. Similar to BSMV, this virus is a positive-strand tripartite RNA virus comprising three genomic RNAs referred to as RNA1, RNA2 and RNA3 (Ahlquist et al. 1984). BMV spread has been recorded in many countries around the world and the virus exhibits one of the broadest host ranges for monocotyledonous plants among viruses and infects, for example, tall fescue (*Festuca arundinacea* Schreb.), barley, rice, maize and legumes (De Jong and Ahlquist 1995). Naturally-occurring BMV variants can differ in their host specificity and, as shown for infectivity of rice, factors on RNA1 and RNA2 can determine the host range for different virus strains (Ding et al. 2006). However, all three genomic RNAs contribute to the systemic infection levels in cowpea (De Jong and Ahlquist 1995).

Previous studies have shown that BMV represents a suitable vector for transient gene expression in barley protoplasts (French et al. 1986). Ding et al. (2006) used a hybrid virus (H-BMV) for infection and silencing of *Pds* and the rice gene *actin1* and observed photobleaching and stunted growth phenotypes, respectively. The hybrid virus was made up of RNA1 and RNA2 from a virus purified from *Festuca arundinacea* Schreb. (F-BMV), which also infected multiple rice accessions, and RNA3 derived from a Russian variant (R-BMV), that failed to infect rice. Although this hybrid virus triggered a VIGS response, the virus symptoms in control plants carrying a hybrid virus without a silencing trigger were much more pronounced if compared to the original F-BMV variant and resembled the virus symptoms observed for the R-BMV variant. A chimeric version, referred to as C-BMV_{A/G}, that utilises RNA1 and RNA2 from F-BMV and a recombinant RNA3 from F-BMV containing a modified intergenic region of R-RNA3 associated with higher RNA expression (Hema and Kao 2004), led to mild virus symptoms, comparable to those observed with F-BMV, but with improved RNA accumulation. In plants silenced for the rubisco activase, the RNA levels of the corresponding gene were reduced by up to 94% compared to control plants. Furthermore, by using crude extract from BMV-infected barley plants to inoculate rice, the same authors demonstrated that the virus can be transmitted mechanically.

13.3.2 Biolistic Approaches

In addition to VIGS, transient particle bombardment based systems have been described for studying gene function in monocotyledonous plants, mainly barley and wheat, that are either based on expression or silencing of candidate genes in single cell assays (Nielsen et al. 1999; Schweizer et al. 2000). Both assays are often restricted to the upper cellular layers of dissected grains, coleoptiles, young roots, or young leaves. To study, for example, the function of barley resistance genes towards *Blumeria graminis* f.sp. *hordei* (*Bgh*) expression of

wild-type *Mlo* was used to mediate susceptibility at the single-cell epidermis level in otherwise *mlo*-resistant plants. Co-expressed target genes are tested by assessing their ability to restore resistance and compromise development of fungal colonies (Shirasu et al. 1999). Using this system, the function and requirement of the race-specific disease resistance gene *Mla6* has been demonstrated in barley (Halterman et al. 2001). In the latter, wild type *Mlo* and green fluorescent protein (GFP) were co-bombarded with functional *Mla*-candidates and challenged with a mildew strain carrying the cognate *Avr* gene. Cells, which have been successfully reverted to the resistant phenotype, prevented GFP expression as programmed cell death occurred following race-specific recognition. Similarly, the function of the resistance candidates genes *HvMla1*, *HvMla7*, *HvMla10*, *HvMla12* and *HvMla13*, identified through map-based cloning approaches, has been successfully demonstrated with variations of the aforementioned assay (Douchkov et al. 2005). However, one drawback of this system is the dependency on full-length cDNA, which prevents high throughput analysis of, for example, cDNA libraries.

The second approach uses double-stranded RNA (dsRNA) for sequence-specific gene silencing and has initially been tested in barley for silencing co-expressed exogenous genes, such as GUS and GFP, but also endogenous sequences, such as dihydroflavonol-4-reductase, involved in anthocyanin biosynthesis (Schweizer et al. 2000). This biolistic approach to transiently induce gene silencing is termed TIGS and has enabled the functional confirmation of major defence-related genes such as *HvSgt1* and *HvRar1*, which are required for many *R* gene-dependent resistance mechanisms (Azevedo et al. 2002). Recently, by utilizing GATEWAYTM based cloning technology, Douchkov et al. (2005) developed a high throughput system to characterize genes required for *R* gene and *mlo* mediated resistance as well as basal defence and non-host resistance. Following TIGS of candidate genes in barley epidermal cells and subsequent infection with *Bgh*, the formation of haustoria as a marker of a compatible interaction in silenced cells was used as a measurement of resistance or susceptibility. This system, in combination with a motorized microscope that automatically evaluates specimens, represents a powerful tool to identify candidate genes that potentially influence the compatible or incompatible outcome of an interaction between barley and *Bgh* (Ihlow et al. 2008).

13.3.3 Antisense Oligodeoxynucleotide

Recently, the use of antisense oligodeoxynucleotides (ODN) as sequence-specific inhibitors of gene expression, and hence as a tool to infer gene function, has emerged in barley and has the potential to evolve as a cheap and truly high-throughput method for functional genomics studies in plants (Sun et al. 2005, 2007). Discovered almost 30 years ago (reviewed in Goodchild 2004), the potential of antisense ODNs, not only as molecules that interfere with gene expression but

also as therapeutics for a wide variety of diseases including viral infections and cancer, was immediately recognized. Indeed, in 1998 the first antisense oligo drug against CMV retinitis, Vitravene, was approved (Marwick 1998). The technology is now widely used in animal science and cancer research (reviewed in Tomita et al. 2003) but has also been reported in plants, first in cell suspension cultures (Tsutsumi et al. 1992), then in cell wall deficient structures such as pollen tubes (Moutinho et al. 2001) and finally in entire barley leaves (Sun et al. 2005, 2007).

Typically, antisense ODN comprise small, single-stranded nucleotides ranging from 12 to 25 nt in length that are specifically designed to hybridise *in vivo* and/or *in vitro* to cognate mRNA in a sequence-specific manner (Sun et al. 2005). The mechanism of how antisense ODN molecules interfere with RNA expression is not fully understood but two main mechanisms, steric hindrance of splicing or the mRNA translational machinery and degradation of mRNA via an RNase H-dependent pathway, have been shown to be involved (Dias and Stein 2002).

Despite the proposed simple mode of action, the choice of sequence and chemistry is crucial in the design of antisense ODNs (Dias and Stein 2002). Mainly due to 5' and 3' endonuclease activities, the half life of standard phosphodiester oligodeoxynucleotides within cells is limited and can, in some extreme cases, be less than a few minutes (Dagle et al. 1991). Numerous modifications of nucleotides have been described that enhance the stability of ODNs in cells but these can either be toxic to cell tissue, result in loss of RNase H-dependent RNA degradation, or result in reduced uptake [reviewed in Dias and Stein (2002)]. So far, the most promising results have been obtained with a high affinity DNA analogue, locked nucleic acid (LNA), which comprises a bicyclic RNA analogue where the ribose sugar is structurally constrained by a methylene bridge between the 2'-oxygen and the 4'-carbon atoms (Wahlestedt et al. 2000).

In addition to the chemistry of antisense ODNs, the conformation and thermodynamic properties of the target mRNA requires thorough analysis as secondary and tertiary folding structures can render mRNA inaccessible to ODNs (Ding and Lawrence 2001). Although predicting secondary structures of the target RNA to aid the identification of effective target sequences can somewhat improve the success rate, the rate of non-effective ODNs remains relatively high (over 50%). The mechanism of naked oligonucleotide uptake into mammalian and plant systems is still poorly understood. Interestingly, when ODN uptake was shown for the first time into barley leaf cells, petioles were submerged in a sucrose solution and solutions lacking sucrose failed to deliver ODNs (Sun et al. 2005, 2007). By using different sugars Sun et al. (2007) could demonstrate that only those sugars that are metabolized, including glucose, fructose, sucrose and maltose, promote ODN uptake in a concentration-dependent manner. Further experiments suggested that the ODNs are translocated via active, H^+ -coupled monosaccharide transporters and/or sucrose translocators and that sucrose also positively influences uptake of ODNs into human HeLa cells. ODN uptake and subsequent inhibition of the

transcription factor gene *SUSIBA2* in barley leaves and endosperm of barley spike cultures (Sun et al. 2007) revealed a block of downstream effects on sugar signalling, including downregulation of *Iso1* and *SBEIIB* mRNA and resulted in altered starch synthesis.

13.4 Stable Genetic Transformation

The stable integration of recombinant DNA into the plant genome is required to manipulate transgene expression, and thereby facilitate the analysis of gene function throughout development. Stable transformation also allows the consequences of a given transformation event to be studied over several generations, and under various environmental conditions. It also avoids any possibility that the genetic effects of the transgene are mistaken for those which arise from the stress of the transformation process itself, and allows a proper correlation to be established between the transgenic phenotype and either transcript or protein abundance. Experimental data derived from stable transgenic plants are therefore inherently more reliable than those based on transient expression experiments.

Unlike an analysis based on a loss-of-function mutant, an integrated transgene can deliver an increased level of functional expression, which is itself potentially informative. Furthermore, anti-sense and inverted repeat constructs typically result not in the complete loss, but rather in a quantitative reduction in the expression of a target gene; thus where loss-of-function mutants are nonviable, transgenics give the opportunity to study gene function via the attenuation of gene expression. Sets of independent transgenic lines showing a range of target gene expression level can be used to correlate gene expression with phenotype, and this is capable of producing a detailed picture of gene function and regulation. In order to minimize the risk of pleiotropic effects, a specific promoter can be combined with a transgene, to ensure its correct temporal and spatial expression, thereby avoiding its expression in either non-target cell types, or during non-target stages of development. Below, we review current opportunities to generate and deploy stable transgenesis in the temperate small grain *Triticeae* cereals.

13.4.1 Transfer of Recombinant DNA into Plant Cells

A prerequisite for successful transgenesis is that following the introduction of the transgene, the transformed cell can proliferate and be regenerated into a viable plant. Unlike the majority of dicotyledonous species, the *Triticeae* species typically cannot be regenerated *in vitro* from leaf or root tissue, and are therefore considered to be particularly challenging in the context of whole plant regeneration. The most important agents of plant transformation, the soilborne

Agrobacterium tumefaciens, evolved their infection and gene transfer machinery (which includes both the suppression and the recruitment of specific host factors; Veena et al. 2003) as an interaction with dicotyledonous plants. As a result, it is scarcely surprising that *Agrobacterium*-mediated gene transfer works so poorly in the *Triticeae* species, and that so much research effort has been expended towards developing other methods of gene transfer.

The earliest gene transfers in the *Triticeae* were confined to transient expression. Once stable transformation of cell lines was achieved via bombardment with DNA-coated metal particles (referred to as biolistic gene transfer), regeneration was the next hurdle to be overcome. Vasil et al. (1992) were the first to generate stably transgenic wheat plants by targeting not cell lines, but embryogenic callus. This success encouraged the development of a number of technical improvements to the biolistic gene transfer procedure and to whole plant regeneration from such explants (Weeks et al. 1993). Related protocols were then rapidly developed for barley (Wan and Lemaux 1994), triticale (Zimny et al. 1995), durum wheat (Bommineni et al. 1997) and rye (Castillo et al. 1994). Alternative recipient tissues were identified, such as the immature inflorescence (Barcelo et al. 1994). In addition, approaches were developed to induce isolated barley microspores (Jaehne et al. 1994) and barley shoot meristematic cultures (Zhang et al. 1999) to undergo embryogenesis and thus become suitable for biolistic gene transfer and the production of stable transgenic plants.

Stable *Agrobacterium*-mediated transformation of *Triticeae* cereals has only been achieved fairly recently. Cheng et al. (1997) and Tingay et al. (1997) were the first to successfully and reproducibly regenerate stable transformants by this route, using immature embryos of wheat and barley, respectively. Since this time, the efficiency of transformation has risen to at least 10% (ten independent primary transgenic plants per 100 inoculated embryos) in barley (Matthews et al. 2001; Hensel et al. 2008), a level far greater than that attainable using biolistic gene transfer. Recent protocols for wheat and rye claim a transformation efficiency of the order of 5% (Wu et al. 2003; Hu et al. 2003; Popelka and Altpeter 2003), equivalent to the level routinely achieved by biolistic gene transfer (Rasco-Gaunt et al. 2001; Popelka et al. 2003). Plant regeneration and transformation efficiencies are, however, highly genotype-dependent. Barley cv. 'Golden Promise' and various selections of wheat cv. 'Bobwhite' are well established as transformation-efficient genotypes. It can often be experimentally desirable, however, to transform a genotype having specific properties, for example a particular field trait or an extreme reaction to pathogen infection. As a result, the range of genotypes able to be transformed has been extended over the recent years in both wheat (Rasco-Gaunt et al. 2001; Wu et al. 2003) and barley (Wang et al. 2001; Murray et al. 2004; Kumlehn et al. 2006; Hensel et al. 2008). Generally, however, transformation efficiencies remain lower than in the standard lines.

Alternative recipient cells and tissues compatible with *Agrobacterium*-based transformation have been identified. While isolated immature pollen and pollen-derived callus of barley were able to generate transgenic cell lines (Wu et al. 1998), a recent report has described a reproducible method for *Agrobacterium*-mediated gene transfer of embryogenic pollen cultures (Kumlehn et al. 2006). This has been further developed to allow haploid transgenic individuals to be selected at the plantlet stage by flow-cytometry, whereafter they can be subjected to diploidization via colchicine treatment (J. Kumlehn et al. unpublished data). The major advantage of this procedure is that its product is a population of homozygous primary transgenic plants, thereby avoiding the self-fertilization steps normally necessary to fix the transgene in the T2 generation, and to validate homozygosity through segregation analysis in T3. Thus, plants can be used for the collection of consistent phenotypic and gene expression data already in the T1 generation. Pursuing a different approach, Holme et al. (2006) were able to infect isolated ovules from barley with *Agrobacterium*, and the major potential of this approach lies in the lack of dependence on the genotype of donor for the ability to regenerate transgenic plants. A further advantage is that this method is still reasonably efficient without the use of a plant selectable marker.

13.4.2 Patterns of DNA-Integration

An analysis of biolistic gene transfer events in wheat has revealed that transgene sequences are integrated at various genomic sites, that there is little evidence for any induction of chromosomal rearrangement, and that transgene expression level is more strongly associated with the choice of promoter sequence than with the location of the integration site (Jackson et al. 2001). In a detailed description of the patterns and sites of T-DNA integration in transgenic barley, Stahl et al. (2002) showed that these are similar to what has been observed in dicotyledonous recipients. About 50% of the events represented the insertion of multiple T-DNAs at a single locus, and all these events consisted of a tandem arrangement of the transgene (no head-to-head or tail-to-tail orientations). About one third of the T-DNA/plant DNA junctions were found in actively transcribed retrotransposon sequences, which is hardly surprising, given the high copy number of transposon sequences in the barley genome. Another study revealed that the right T-DNA ends were largely conserved, whereas the left ends were more variable (Fang et al. 2002). Integration sites in barley, following either biolistic or *Agrobacterium*-based gene transfer to immature embryos, were physically located by fluorescence *in situ* hybridisation (FISH) by Salvo-Garrido et al. (2004). This analysis resulted in the identification of 23 integration sites spread over five of the seven barley chromosomes. The pattern of integration was non-randomly distributed, with a bias towards the telomeric and subtelomeric regions.

13.4.3 The Design of Transformation Vectors

The identification and deployment of appropriate promoter sequences to drive the expression of a transgene is important for both the development of the transformation technology itself, and for its application as a tool in basic and applied research. While promoters associated with constitutive expression have been much employed, along with selectable marker and reporter genes, for establishing transformation technology, it is increasingly desirable to turn to cell-, tissue- or organ-specific ones, as well as to those which can be exogenously induced. Most dicotyledonous plant promoters, covering a wide range of specificities and expression strengths, are barely, if at all, functional in a monocotyledonous recipient. Therefore, the choice of promoters available for the *Triticeae* cereals remains rather limited.

The most commonly used constitutive promoters have been maize ubiquitin 1 (Christensen and Quail 1996; Furtado and Henry 2005), cauliflower mosaic virus 35S (Furtado and Henry 2005) and rice actin 1 (Vickers et al. 2006). Of the few specific promoters available, most drive gene expression in the endosperm; these include promoters for wheat low molecular weight glutenin D1 (Stoeger et al. 1999), wheat high molecular weight glutenin subunits 1D-1 or 1Ax1 (Lamacchia et al. 2001; Yao et al. 2006), wheat ADP-glucose pyrophosphorylase (Thornycroft et al. 2003), wheat granule-bound starch synthase 1 (Kluth et al. 2002), durum wheat puroindolines *pin A* or *pin B* (Wiley et al. 2007), barley hordein B1 or D (Cho et al. 1999, 2002; Vickers et al. 2006), oat globulin 1 (Vickers et al. 2006) and rice glutelin B-1 (Patel et al. 2000). Other functional *Triticeae* promoters have been derived from the nucellar projection-specific *jeekyll* gene of barley (Radchuk et al. 2006), from barley embryo-specific *Asi* and *Em* (Furtado et al. 2003; Furtado and Henry 2005), from barley *Lem1* and *Lem2* (Skadsen et al. 2002; Tilahun et al. 2006), and from the epidermis-specific glutathione-S-transferase A1 (*GstA1*) of wheat (Altpeter et al. 2005).

The configuration of the expression cassette is responsible for the effect on phenotype, if any, of the transgene, and this in turn allows for some deductions to be made regarding gene function. For instance, the integration of a recombinant copy of an endogenous gene sequence under the control of a constitutive promoter can be expected to deliver an elevated level of expression of the gene, not merely in its normal cellular context, but also in tissues and during developmental windows when the endogenous gene is inactive. By expressing a transgene in a loss-of-function mutant, predicted gene function can be verified through functional complementation (Stein et al. 2005). Alternatively, the integration of a transgene in the anti-sense orientation or in an inverted repeat configuration can often induce a reduction in the target gene expression, in a phenomenon known as RNA interference (RNAi). The mechanism of RNAi relies on the formation of a hairpin loop in the transgene message, and this stretch of double-stranded mRNA is recognized and degraded by the host's post-transcriptional gene silencing machinery. This sequence-specific RNA

degradation applies equally to endogenous gene transcripts, and as a result, their expression is reduced or even abolished.

Transgenes made up of inverted repeats are considered more effective in gene silencing than anti-sense constructs. RNAi can in some cases result in the loss of expression of all members of a gene family, a possibility which is especially useful for functional studies in polyploids. A prominent example is hexaploid wheat, where many genes are represented by three fully functional homoeologous copies. The potential of RNAi as a tool in wheat functional genomics approaches has been recently demonstrated by Travella et al. (2006), who used hairpin transgenes based on sequences of a phytoene desaturase (*Pds*) and the ethylene insensitive 2 (*Ein2*) gene to generate clear phenotypes caused by RNAi-induced down-regulation of all three wheat homoeologs in equal measure. The severity of the transgenic phenotype was clearly associated with both the transcript level of the targeted genes and with the abundance of small RNAs, derived from the gene-specific mRNA degradation process. This particular RNAi-mediated gene silencing proved to be fully heritable.

The majority of the binary vector systems developed to date have been targeted at the transformation of dicotyledonous plants (e.g. Wesley et al. 2001; Curtis and Grossniklaus 2003). Typically, however, these are not useful for the monocotyledons, mainly because of the limited functionality of the plant promoters used to drive either the transgene or the selection marker. Other specific vector elements, in particular the selectable marker and the origin of replication, can also reduce the efficacy of a binary vector. Standard cloning procedures to engineer transformation vectors for over-expression or RNAi-mediated knock-down remain laborious and time-consuming. Recently, however, a set of generic binary vectors has been designed for the stable transformation of monocotyledonous species (Himmelbach et al. 2007). Their modular configuration is intended to simplify the insertion into both over-expression and knock-down vectors of both promoter and effector sequences, and the selective marker. The insertion of the effector sequences is achieved using the GATEWAYTM recombination system. The spectrum of applications has been extended by enabling the pre-testing of constructs in transient expression assays, and by allowing for the transformation of both mono- and dicotyledonous hosts with the same binary vector. The vector set is built from derivatives of the three common constitutive promoters – maize ubiquitin (Furtado and Henry 2005), the doubled enhanced CaMV 35S (Furtado and Henry 2005) and rice actin 1 (McElroy et al. 1990; Vickers et al. 2006). Other derivatives harbouring the wheat glutathione-S transferase promoter (Altpeter et al. 2005) permit the constitutive expression of transgenes in exclusively the leaf epidermis. The functionality of the cereal vector set has been demonstrated in barley by both transient and stable transformation experiments, involving both over-expression and targeted gene silencing.

13.4.4 Insertional Mutagenesis

Insertional mutagenesis has been used extensively in a number of plant species to create representative populations of knock-out mutant lines which are highly valuable for the tagging and the functional characterisation of the interrupted genes. The establishment of such populations can be achieved by the random integration of T-DNAs via *Agrobacterium*-mediated transformation, or by the introduction and remobilisation of appropriately designed transposable elements. While no attempts to produce T-DNA insertion lines have yet been published for *Triticeae* species, a population of transposon insertion lines of barley based on the maize activator/dissociation (*Ac/Ds*) transposon system has been generated by Koprek et al. (2000). A similar population of several hundred independent *Ds* plants has also been generated (Ayliffe et al. 2007). Transposon-based insertional mutagenesis has certain advantages as compared to other approaches to the determination of gene function. For instance, only a few initial transformants are required to generate large numbers of plants carrying transposed elements at different locations. Moreover, transposable elements can be mobilized or immobilized on demand, and transposons were shown to preferentially integrate into genic regions, which is important in large genomes such as those of *Triticeae* species. More recently, Singh et al. (2006) demonstrated that gene tagging is facilitated by high-frequency remobilization of the *Ds* element over multiple generations in barley. They presented data on the remobilisation frequencies of transposants with intact terminal inverted repeats. Moreover, molecular and flanking sequence characterisation of 50 transposant lines was reported and the preferential transposition into genic regions validated. Putative genes tagged in this study encoded cell wall-associated kinases, ABC transporter, ubiquitin conjugating enzyme, terpene synthase, F-Box protein, Mla, and cytochrome P450. Furthermore, it was shown that the frequencies of re-activation are adequate to employ the *Ac/Ds* transposon system in future approaches to the saturation mutagenesis in barley.

13.4.5 Linking Manipulated Gene Expression with Gene Function

Transgenesis plays a key role in the process of validating gene function because it allows specific knock-in or knock-down of the expression of a target gene. The resulting transgenic plants can be compared directly with wild type plants and from this comparison an assessment made about the biological role of the target gene. For example, plant development and architecture are two of the major determinants of a crop's field performance. Many attempts have been made to understand how these are governed at the molecular level. The *Q* (speltoid) gene of wheat, which determines the shape of the ear, is one of the major players in the crop's domestication. Quantitative RT-PCR analysis of transgenic plants has shown that the formation of a speltoid ear is associated

with the level of *Apetala 2*-like (*Wap2*) transcript level (Simons et al. 2006). The reduced *Wap2* transcript level in transgenic speltoid lines was assumed to have been obtained by co-suppression, a rather common phenomenon among the products of biolistic gene transfer. Using RNAi approaches based on einkorn (*Triticum monococcum*) trigger sequences, it was possible to demonstrate that floral initiation in winter wheat is regulated by *Vrn1* and *Vrn2*, which encode, respectively, a MADS-box transcription factor and a zinc-finger CCT domain protein (Yan et al. 2004; Loukoianov et al. 2005).

Since the grain represents the most important product of the cereal plant, much research activity has focused on grain yield and grain quality. Verification of the function of key candidate genes involved in wheat starch biosynthesis (GBSSI – granule-bound starch synthase 1 – and SBEII – starch branching enzyme) was achieved by RNAi technology (Li et al. 2005; Regina et al. 2006). The high molecular weight (HMW) gluten proteins are particularly significant determinants of the bread-making quality of wheat, and their role has been validated by a number of transgenic experiments involving the over-expression or post-transcriptional silencing of the endogenous genes (Altpeter et al. 1996; Blechl and Anderson 1996; Barro et al. 1997; Alvarez et al. 2000). Via its over-expression, the puroindoline gene *pinB* has been shown to play a critical role in the determination of grain hardness – a feature of central importance to the milling and baking industries (Beecher et al. 2002). Also, the regulatory role of transcription factors in grain development has been established by transgene driven over-expression strategies. As an example, the barley transcription factor GAMYB was shown to regulate the expression of hydrolytic enzymes in the aleurone (Murray et al. 2006). Similarly, Radchuk et al. (2006) were able to determine the specific function, within the process of grain development, of the barley *jeekyll* gene.

Disease is a major concern for all crop species and attempts have been made to elucidate resistance gene functions by the use of transgenesis. A functional analysis of the genes implicated in the complex plant-pathogen interaction can help to understand the molecular events occurring during pathogen attack, and this knowledge can be used to derive transgenic strategies for disease resistance. Verification of the function of the barley stem rust resistance gene *Rpg1* cloned by Brueggeman et al. (2002) was achieved by its stable over-expression (Horvath et al. 2003). A crucial role for RAR1 in the control of R protein homeostasis was revealed through the analysis of MLA1, MLA6, and MLA chimeric protein expressed in transgenic barley lines that exhibit a differential RAR1 requirement for resistance to powdery mildew (Bieri et al. 2004). In an attempt to elucidate the function of a small GTPase in the context of infection by *Blumeria graminis f.sp. hordei*, stable over-expression of constitutively activated RACB rendered barley more susceptible to both powdery mildew and abiotic stress (Schultheiss et al. 2005). An enhanced level of the plant's resistance to powdery mildew was found by Altpeter et al. (2005) who ectopically expressed a peroxidase (*Pero*) gene in the wheat leaf epidermis.

Gene function has in some cases been validated by heterologous transgene expression. For instance, a barley apoplastic ribosome inactivation protein expressed in transgenic wheat was shown by Bieri et al. (2000) to confer a degree of resistance against wheat powdery mildew. In another study, transgenic wheat expressing a barley seed class II chitinase gene showed increased resistance against powdery mildew (Bliffeld et al. 1999). Transgenesis has also been exploited to determine the function of host genes involved in the interaction between a cereal plant and a pathogenic virus. Specifically, the barley eukaryotic translation initiation factor 4E (*Eif4E*) was putatively associated with resistance to Bymoviruses on the basis of positional cloning, and when the *Eif4E* allele derived from a susceptible variety was introduced into a resistant genetic background, the resulting change in phenotype provided compelling evidence that the gene represents an essential host factor (Stein et al. 2005).

13.5 Final Remarks

Eighty years after the potential of mutagenesis in plants was first described, the reverse genetics methodology called TILLING has boosted *Triticeae* (and other crop plant) functional genomics. However despite the history of mutagenesis, TILLING was already preceded by transgenic biology, by VIGS and the majority of the transient transformation approaches summarized above. Hopefully, from our discussion, it is clear to the reader that each approach has unique properties and that choosing how and when one rather than an alternative should be adopted depends ultimately upon to the objectives of the experimental situation. Furthermore, while the products of some functional validation strategies are potentially appropriate for deployment in agriculture (an outcome influenced by both political and biological reasons), all are both applicable and effective for the functional characterization of candidate genes.

References

- Ahlquist, P., Dasgupta, R. and Kaesberg, P. (1984) Nucleotide sequence of the brome mosaic virus genome and its implications for viral replication. *J. Mol. Biol.* 172, 369–383.
- Altpeter, F., Vasil, V., Srivastava, V. and Vasil, I. K. (1996) Integration and expression of the high-molecular-weight glutenin subunit 1Ax1 gene into wheat. *Nat. Biotech.* 14, 1155–1159.
- Altpeter, F., Varshney, A., Abderhalden, O., Douchkov, D., Sautter, C., Kumlehn, J., Dudler, R. and Schweizer, P. (2005) Stable expression of a defense-related gene in wheat epidermis under transcriptional control of a novel promoter confers pathogen resistance. *Plant Mol. Biol.* 57, 271–283.
- Alvarez, M. L., Guelman, S., Halford, N. G., Lustig, S., Reggiardo, M. I., Rybabushkina, N., Shewry, P. R., Stein, J. and Vallejos, R. H. (2000) Silencing of HMW glutenins in transgenic wheat expressing extra HMW subunits. *Theor. Appl. Genet.* 100, 319–327.
- Ashburner, M. (1990) *Drosophila: A Laboratory Handbook*. Cold Spring Harbor, NY, USA.

- Ayliffe, M. A., Pallota, M., Langridge, P. and Pryor, A. J. (2007) A barley activation tagging system. *Plant Mol. Biol.* 64, 329–347.
- Azevedo, C., Sadanandom, A., Kitagawa, K., Freialdenhoven, A., Shirasu, K. and Schulze-Lefert, P. (2002) The *RAR1* interactor *SGT1*, an essential component of R gene-triggered disease resistance. *Science* 295, 2073–2076.
- Barcelo, P., Hagel, C., Becker, D., Martin, A. and Lorz, H. (1994) Transgenic cereal (*Tritordeum*) plants obtained at high efficiency by microprojectile bombardment of inflorescence tissue. *The Plant J.* 5, 583–592.
- Barro, F., Rooke, L., Békés, F., Gras, P., Tatham, A. S., Fido, R. J., Lazzeri, P., Shewry, P. R. and Barcelo, P. (1997) Transformation of wheat with HMW subunit genes results in improved functional properties. *Nat. Biotech.* 15, 1295–1299.
- Beecher, B., Bettge, A., Smidansky, E. and Giroux, M. J. (2002) Expression of wild-type *pinB* sequence in transgenic wheat complements a hard phenotype. *Theor. Appl. Genet.* 105, 870–877.
- Bieri, S., Potrykus, I. and Futterer, J. (2000) Expression of active barley seed ribosome-inactivating protein in transgenic wheat. *Theor. Appl. Genet.* 100, 755–763.
- Bieri, S., Mauch, S., Shen, Q. H., Peart, H., Devoto, A., Casais, C., Ceron, F., Schulze, S., Steinbüß, H. H., Shirasu, K. and Schulze-Lefert, P. (2004) *RAR1* positively controls steady state levels of barley *MLA* resistance proteins and enables sufficient *MLA6* accumulation for effective resistance. *The Plant Cell* 16, 3480–3495.
- Blechl, A. E. and Anderson, O. D. (1996) Expression of a novel high-molecular-weight glutenin subunit gene in transgenic wheat. *Nat. Biotech.* 14, 875–897.
- Bliffeld, M., Mundy, J., Potrykus, I. and Futterer, J. (1999) Genetic engineering of wheat for increased resistance to powdery mildew disease. *Theor. Appl. Genet.* 98, 1079–1086.
- Bommineni, V. R., Jauhar, P. P. and Peterson, T. S. (1997) Transgenic durum wheat by microprojectile bombardment of isolated scutella. *J. Hered.* 88, 475–481.
- Brigneti, G., Martin-Hernández, A. M., Jin, H., Chen, J., Baulcombe, D. C., Baker, B. and Jones, J. D. G. (2004) Virus-induced gene silencing in *Solanum* species. *The Plant J.* 39, 264–272.
- Brueggeman, R., Rostoks, N., Kudrna, D., Kilian, A., Han, F., Chen, J., Druka, A., Steffensen, B. and Kleinhofs, A. (2002) The barley stem rust resistance gene *Rpg1* is a novel disease-resistance gene with homology to receptor kinases. *Proc. Natl. Acad. Sci. USA* 99, 9328–9333.
- Bruun-Rasmussen, M., Madsen, C. T., Jessing, S. and Albrechtsen, M. (2007) Stability of Barley stripe mosaic virus-induced gene silencing in barley. *Mol. Plant-Microbe Interact.* 20, 1323–1331.
- Büschges, R., Hollricher, K., Panstruga, R., Simons, G., Wolter, M., Frijters, A., van Daelen, R., van der Lee, T., Diergaarde, P., Groenendijk, J., Töpsch, S., Vos, P., Salamini, F. and Schulze-Lefert, P. (1997) The barley *Mlo* gene: A novel control element of plant pathogen resistance. *Cell* 88, 695–705.
- Caldwell, D. G., McCallum, N., Shaw, P., Muehlbauer, G. J., Marshall, D. F. and Waugh, R. (2004) A structured mutant population for forward and reverse genetics in barley (*Hordeum vulgare* L.). *The Plant J.* 40, 143–150.
- Castillo, A., Vasil, V. and Vasil, I. K. (1994) Rapid production of fertile transgenic plants of rye (*Secale cereale* L.). *Bio/Technol.* 12, 1366–1371.
- Cheng, M., Fry, J. E., Pang, S., Zhou, H., Hironaka, C., Duncan, D. R., Conner, T. W. and Wan, Y. (1997) Genetic transformation of wheat mediated by *Agrobacterium tumefaciens*. *Plant Physiol.* 115, 971–980.
- Cho, M. J., Choi, H. W., Buchanan, B. B. and Lemaux, P. G. (1999) Inheritance of tissue specific expression of barley hordein promoter-*uidA* fusions in transgenic barley plants. *Theor. Appl. Genet.* 98, 1253–1262.
- Cho, M. J., Choi, H. W., Jiang, W., Ha, C. D. and Lemaux, P. G. (2002) Endosperm-specific expression of green fluorescent protein driven by the hordein promoter is stably inherited in transgenic barley (*Hordeum vulgare*) plants. *Physiol. Plantarum* 115, 144–151.

- Christensen, A. H. and Quail, P. H. (1996) Ubiquitin promoter-based vectors for high level expression of selectable and/or screenable marker genes in monocotyledonous plants. *Trans. Res.* 5, 213–218.
- Curtis, M. D. and Grossniklaus, U. (2003) A GATEWAY cloning vector set for high-throughput functional analysis of genes *in planta*. *Plant Physiol.* 133, 462–469.
- Dagle, J. M., Weeks, D. L. and Walder, J. A. (1991) Pathways of degradation and mechanism of action of antisense oligonucleotides in *Xenopus laevis* embryos. *Antisense Res. Dev.* 1, 11–20.
- Dalmaï, M., Schmidt, J., Le Signor, C., Moussy, F., Burstin, J., Savoie, V., Aubert, G., Brunaud, V., de Oliveira, Y., Guichard, C., Thompson, R. and Bendahmane, A. (2008) UTILdb, a *Pisum sativum* *in silico* forward and reverse genetics tool. *Genome Biol.* 9, R43.
- De Jong, W. and Ahlquist, P. (1995) Host-specific alterations in viral RNA accumulation and infection spread in a brome mosaic virus isolate with an expanded host range. *J. Virol.* 69, 1485–1492.
- Dias, N. and Stein, C. A. (2002) Antisense Oligonucleotides: Basic Concepts and Mechanisms. *Mol. Cancer Ther.* 1, 347–355.
- Ding, X. S., Schneider, W. L., Chaluvadi, S. R., Mian, M. A. R. and Nelson, R. S. (2006) Characterization of a Brome mosaic virus strain and its use as a vector for gene silencing in monocotyledonous hosts. *Mol. Plant-Microbe Interact.* 19, 1229–1239.
- Ding, Y. and Lawrence, C. E. (2001) Statistical prediction of single-stranded regions in RNA secondary structure and application to predicting effective antisense target sites and beyond. *Nucleic Acids Res.* 29, 1034–1046.
- Douchkov, D., Nowara, D., Zierold, U. and Schweizer, P. (2005) A high-throughput gene-silencing system for the functional assessment of defense-related genes in barley epidermal cells. *Mol. Plant-Microbe Interact.* 18, 755–761.
- Edwards, M. C. (1995) Mapping of the seed transmission determinants of barley stripe mosaic virus. *Mol. Plant-Microbe Interact.* 8, 906–915.
- Fang, Y. D., Akula, C. and Altpeter, F. (2002) *Agrobacterium*-mediated barley (*Hordeum vulgare* L.) transformation using green fluorescent protein as a visual marker and sequence analysis of the T-DNA:genomic DNA junctions. *J. Plant Physiol.* 159, 1131–1138.
- French, R., Janda, M. and Ahlquist, P. (1986) Bacterial gene inserted in an engineered RNA virus: efficient expression in monocotyledonous plant cells. *Science* 231, 1294–1297.
- Furtado, A., Henry, R., Scott, K. and Meech, S. (2003) The promoter of the *asi* gene directs expression in the maternal tissues of the seed in transgenic barley. *Plant Mol. Biol.* 52, 787–799.
- Furtado, A. and Henry, R. J. (2005) The wheat Em promoter drives reporter gene expression in embryo and aleurone tissue of transgenic barley and rice. *Plant Biotech. J.* 3, 421–434.
- Goodchild, J. (2004) Oligonucleotide therapeutics: 25 years agrowing. *Curr. Opin. Mol. Ther.* 6, 120–128.
- Greene, E. A., Codomo, C. A., Taylor, N. E., Henikoff, J. G., Till, B. J., Reynolds, S. H., Enns, L. C., Burtner, C., Johnson, J. E., Odden, A. R., Comai, L. and Henikoff, S. (2003) Spectrum of chemically induced mutations from a large-scale reverse-genetic screen in *Arabidopsis*. *Genetics* 164, 731–740.
- Halterman, D., Zhou, F., Wei, F., Wise, R. P. and Schulze-Lefert, P. (2001) The *MLA6* coiled-coil, NBS-LRR protein confers *AvrMla6*-dependent resistance specificity to *Blumeria graminis* f. sp. *hordei* in barley and wheat. *The Plant J.* 25, 335–348.
- Haupt, S., Duncan, G. H., Holzberg, S. and Oparka, K. J. (2001) Evidence for symplastic phloem unloading in sink leaves of barley. *Plant Physiol.* 125, 209–218.
- Hein, I., Barciszewska-Pacak, M., Hrubikova, K., Williamson, S., Dinesen, M., Soenderby, I. E., Sundar, S., Jarmolowski, A., Shirasu, K. and Lacomme, C. (2005) Virus-induced gene silencing-based functional characterization of genes associated with powdery mildew resistance in barley. *Plant Physiol.* 138, 2155–2164.

- Hema, M. and Kao, C. C. (2004) Template sequence near the initiation nucleotide can modulate brome mosaic virus RNA accumulation in plant protoplasts. *J. Virol.* 78, 1169–1180.
- Hensel, G., Valkov, V., Middlefell-Williams, J. and Kumlehn, J. (2008) Efficient generation of transgenic barley: the way forward to modulate plant-microbe interactions. *J. Plant Physiol.* 165, 71–82.
- Himmelbach, A., Zierold, U., Hensel, G., Riechen, J., Douchkov, D., Schweizer, P. and Kumlehn, J. (2007) A set of modular binary vectors for the transformation of cereals. *Plant Physiol.* 145, 1192–1200.
- Hodgdon, A. L., Marcus, A. H., Arenaz, P., Rosichan, J. L., Bogyo, T. P. and Nilan, R. A. (1981) Ontogeny of the barley plant as related to mutation expression and detection of pollen mutations. *Environ. Health Perspect.* 37, 5–7.
- Holme, I. B., Brinch-Pedersen, H., Lange, M. and Holm, P. B. (2006) Transformation of barley (*Hordeum vulgare* L.) by *Agrobacterium tumefaciens* infection of *in vitro* cultured ovules. *Plant Cell Rep.* 25, 1325–1335.
- Holzberg, S., Brosio, P., Gross, C. and Pogue, G. P. (2002) Barley stripe mosaic virus-induced gene silencing in a monocot plant. *The Plant J.* 30, 315–327.
- Horvath, H., Rostoks, N., Brueggeman, R., Steffensen, B., von Wettstein, D. and Kleinhofs, A. (2003) Genetically engineered stem rust resistance in barley using the *Rpg1* gene. *Proc. Natl. Acad. Sci. USA* 100, 364–369.
- Hu, T., Metz, S., Chay, C., Zhou, H. P., Biest, N., Chen, G., Cheng, M., Feng, X., Radionenko, M., Lu, F. and Fry, J. (2003) *Agrobacterium*-mediated large-scale transformation of wheat (*Triticum aestivum* L.) using glyphosate selection. *Plant Cell Rep.* 21, 1010–1019.
- Ihlow, A., Schweizer, P. and Seiffert, U. (2008) A high-throughput screening system for barley/powdery mildew interactions based on automated analysis of light micrographs. *BMC Plant Biol.* 8, 6.
- Jackson, S. A., Zhang, P., Chen, W. P., Phillips, R. L., Friebe, B., Muthukrishnan, S. and Gill, B. S. (2001) High-resolution structural analysis of biolistic transgene integration into the genome of wheat. *Theor. Appl. Genet.* 103, 56–62.
- Jaehne, A., Becker, D., Bretschneider, H. and Loez, H. (1994) Regeneration of transgenic, microspore-derived, fertile barley. *Theor. Appl. Genet.* 89, 525–533.
- Johansen, E., Edwards, M. C. and Hampton, R. O. (1994) Seed transmission of viruses: current perspectives. *Annu. Rev. Phytopathol.* 32, 363–386.
- Johnson, J. A., Bragg, J. N., Lawrence, D. M. and Jackson, A. O. (2003) Sequence elements controlling expression of Barley stripe mosaic virus subgenomic RNAs *in vivo*. *Virology* 313, 66–80.
- Kanyuka, K., Druka, A., Caldwell, D. G., Tymon, A., McCallum, N., Waugh, R. and Adams, M. J. (2005). Evidence that the recessive bymovirus resistance locus *rym4* in barley corresponds to the eukaryotic translation initiation factor 4E gene. *Mol. Plant Path.* 6, 449–458.
- Kluth, A., Sprunck, S., Becker, D., Loez, H. and Luettticke, S. (2002) 5' deletion of a gbss1 promoter region from wheat leads to changes in tissue and developmental specificities. *Plant Mol. Biol.* 49, 669–682.
- Koprek, T., McElroy, D., Louwerse, J., Williams-Carrier, R. and Lemaux, P. G. (2000) An efficient method for dispersing Ds elements in the barley genome as a tool for determining gene function. *The Plant J.* 24, 253–264.
- Kumlehn, J., Serazetdinova, L., Hensel, G., Becker, D. and Loez, H. (2006) Genetic transformation of barley (*Hordeum vulgare* L.) via infection of androgenetic pollen cultures with *Agrobacterium tumefaciens*. *Plant Biotech. J.* 4, 251–261.
- Lacomme, C., Hrubikova, K. and Hein, I. (2003) Enhancement of virus-induced gene silencing through viral-based production of inverted-repeats. *The Plant J.* 34, 543–553.
- Lamacchia, C., Shewry, P. R., Di Fonzo, N., Forsyth, J. L., Harris, N., Lazzeri, P. A., Napier, J. A., Halford, N. G. and Barcelo, P. (2001) Endosperm-specific activity of a storage protein gene promoter in transgenic wheat seed. *J. Exp. Bot.* 52, 243–250.

- Li, S. L. and Redei, G. P. (1969) Estimation of mutation rate in autogamous diploids. *Radiation Botany* 9, 125–131.
- Li, J. R., Zhao, W., Li, Q. Z., Ye, X. G., An, B. Q., Li, X. and Zhang, X. S. (2005) RNA silencing of *Waxy* gene results in low levels of amylose in the seeds of transgenic wheat (*Triticum aestivum* L.). *Acta Genet. Sin.* 32, 846–854.
- Loukoianov, A., Yan, L., Blechl, A., Sanchez, A. and Dubcovsky, J. (2005) Regulation of *VRN-1* vernalization genes in normal and transgenic polyploid wheat. *Plant Physiol.* 138, 2364–2373.
- Marwick, C. (1998) First “antisense” drug will treat CMV retinitis. *J. Am. Med. Assoc.* 280, 871.
- Matthes, M., Singh, R., Cheah, S. C. and Karp, A. (2001) Variation in oil palm (*Elaeis guineensis* Jacq.) tissue culture-derived regenerants revealed by AFLPs with methylation-sensitive enzymes. *Theor. Appl. Genet.* 102, 971–979.
- Matthews, P. R., Wang, M. B., Waterhouse, P. M., Thornton, S., Fieg, S. J., Gubler, F. and Jacobsen, J. V. (2001) Marker gene elimination from transgenic barley, using co-transformation with adjacent ‘twin T-DNAs’ on a standard *Agrobacterium* transformation vector. *Mol. Breed.* 7, 195–202.
- McCallum, C. M., Comai, L., Greene, E. A. and Henikoff, S. (2000a) Targeted screening for induced mutations. *Nat. Biotech.* 18, 455–457.
- McCallum, C. M., Comai, L., Greene, E. A. and Henikoff, S. (2000b) Targeting induced local lesions in genomes (TILLING) for plant functional genomics. *Plant Physiol.* 123, 439–442.
- McElroy, D., Zhang, W. G., Cao, J. and Wu, R. (1990) Isolation of an efficient actin promoter for use in rice transformation. *The Plant Cell* 2, 163–171.
- McKinney, H. H. and Greeley, L. W. (1965) Biological characteristics of barley stripe mosaic virus strains and their evolution. *USDA Tech. Bull.* 1324.
- Moissiard, G. and Voinnet, O. (2006) RNA silencing of host transcripts by cauliflower mosaic virus requires coordinated action of the four *Arabidopsis* Dicer-like proteins. *Proc. Natl. Acad. Sci. USA* 103, 19593–19598.
- Moutinho, A., Camacho, L., Haley, A., Pais, M. S., Trewavas, A. and Malho, R. (2001) Antisense perturbation of protein function in living pollen tubes. *Sex. Plant Reprod.* 14, 101–104.
- Murray, F., Brettell, R., Matthews, P., Bishop, D. and Jacobsen, J. (2004) Comparison of *Agrobacterium*-mediated transformation of four barley cultivars using the GFP and GUS reporter genes. *Plant Cell Rep.* 22, 397–402.
- Murray, F., Matthews, P., Jacobsen, J. and Gubler, F. (2006) Increased expression of HvGAMYB in transgenic barley increases hydrolytic enzyme production by aleurone cells in response to gibberellin. *J. Cereal Sci.* 44, 317–322.
- Nielsen, K., Olsen, O. and Oliver, R. (1999) A transient expression system to assay putative antifungal genes on powdery mildew infected barley leaves. *Physiol. Mol. Plant Pathol.* 54, 1–12.
- Ng, P. C. and Henikoff, S. (2003) SIFT: predicting amino acid changes that affect protein function. *Nucleic Acids Res.* 31, 3812–3814.
- Oikawa, A., Rahman, A., Yamashita, T., Taira, H. and Kidou, S. I. (2007) Virus-induced gene silencing of P23k in barley leaf reveals morphological changes involved in secondary wall formation. *J. Exp. Bot.* 58, 2617–2625.
- Patel, M., Johnson, J. S., Brettell, R. I. S., Jacobsen, J. and Xue, G. P. (2000) Transgenic barley expressing a fungal xylanase gene in the endosperm of the developing grains. *Mol. Breed.* 6, 113–123.
- Perry, J. A., Wang, T. L., Welham, T. J., Gardner, S., Pike, J. M., Yoshida, S. and Parniske, M. (2004) A TILLING reverse genetics tool and a web-accessible collection of mutants of the legume *Lotus japonicus*. *Plant Physiol.* 131, 866–871.
- Popelka, J. C. and Altpeter, F. (2003) *Agrobacterium tumefaciens*-mediated genetic transformation of rye (*Secale cereale* L.). *Mol. Breed.* 11, 203–211.
- Popelka, J. C., Xu, J. and Altpeter, F. (2003) Generation of rye plants with low copy number after biolistic gene transfer and production of instantly marker-free transgenic rye. *Transgenic Res.* 12, 587–596.

- Qui, P., Shandilya, H., D'Alessio, J. M., O'Connor, K., Durocher, J. and Gerard, G. F. (2004) Mutation detection using Surveyor nuclease. *Biotechniques* 36, 702.
- Radchuk, V., Borisjuk, L., Radchuk, R., Steinbiss, H. H., Rolletschek, H., Broeders, S. and Wobus, U. (2006) *Jekyll* encodes a novel protein involved in the sexual reproduction of barley. *The Plant Cell* 18, 1652–1666.
- Rasco-Gaunt, S., Riley, A., Cannell, M., Barcelo, P. and Lazzeri, P. A. (2001) Procedures allowing the transformation of a range of European elite wheat (*Triticum aestivum* L.) varieties via particle bombardment. *J. Exp. Bot.* 52, 865–874.
- Regina, A., Bird, A., Topping, D., Bowden, S., Freeman, J., Barsby, T., Kosar-Hashemi, B., Li, Z., Rahman, S. and Morell, M. (2006) High-amylose wheat generated by RNA interference improves indices of large-bowel health in rats. *Proc. Natl. Acad. Sci. USA* 103, 3546–3551.
- Ruiz, M. T., Voinnet, O. and Baulcombe, D. C. (1998) Initiation and maintenance of virus-induced gene silencing. *The Plant Cell* 10, 937–946.
- Salvo-Garrido, H., Travella, S., Bilham, L. J., Harwood, W. A. and Snape, J. W. (2004) The distribution of transgene insertion sites in barley determined by physical and genetic mapping. *Genetics* 167, 1371–1379.
- Schultheiss, H., Hensel, G., Imani, J., Broeders, S., Sonnewald, U., Kogel, K. H., Kumlehn, J. and Hueckelhoven, R. (2005) Ectopic expression of constitutively activated RACB small GTPase in barley enhances susceptibility to powdery mildew and abiotic stress. *Plant Physiol.* 139, 353–362.
- Schweizer, P., Pokorny, J., Schulze-Lefert, P. and Dudler, R. (2000) Technical advance. Double-stranded RNA interferes with gene function at the single-cell level in cereals. *The Plant J.* 24, 895–903.
- Scofield, S. R., Huang, L., Brandt, A. S. and Gill, B. S. (2005) Development of a virus-induced gene-silencing system for hexaploid wheat and its use in functional analysis of the *Lr21*-mediated leaf rust resistance pathway. *Plant Physiol.* 138, 2165–2173.
- Shirasu, K., Nielsen, K., Piffanelli, P., Oliver, R. and Schulze-Lefert, P. (1999) Cell-autonomous complementation of *mlo* resistance using a biolistic transient expression system. *The Plant J.* 17, 293–299.
- Simons, K. J., Fellers, J. P., Trick, H. N., Zhang, Z., Tai, Y., Gill, B. K. and Faris, J. D. (2006) Molecular characterization of the major wheat domestication gene *Q*. *Genetics* 172, 547–555.
- Singh, J., Zhang, S., Chen, C., Cooper, L., Bregitzer, P., Sturbaum, A., Hayes, P. M. and Lemaux, P. G. (2006) High-frequency Ds remobilization over multiple generations in barley facilitates gene tagging in large genome cereals. *Plant Mol. Biol.* 62, 937–950.
- Skadsen, R. W., Sathish, P., Federico, M. L., Abebe, T., Fu, J. and Kaeppler, H. F. (2002) Cloning of the promotor for a novel barley gene, *Lem1*, and its organ-specific promotion of *Gfp* expression in lemma and palea. *Plant Mol. Biol.* 49, 545–555.
- Slade, A. J., Fuerstenberg, S. I., Loeffler, D., Steine, M. N. and Facciotti, D. (2005) A reverse genetic, non-transgenic approach to wheat crop improvement by TILLING. *Nature Biotechnology* 23, 75–81.
- Stadler, L. J. (1928a) Mutations in barley induced by x-rays and radium. *Science* 68, 186–197.
- Stadler, L. J. (1928b) Genetic effects of X-rays in maize. *Proc. Natl. Acad. Sci. USA* 14, 69–75.
- Stahl, R., Horvath, H., Van Fleet, J., Voetz, M., von Wettstein, D. and Wolf, N. (2002) T-DNA integration into the barley genome from single and double cassette vectors. *Proc. Natl. Acad. Sci. USA* 99, 2146–2151.
- Stein, N., Perovic, D., Kumlehn, J., Pellic, B., Stracke, S., Streng, S., Ordon, F. and Graner, A. (2005) The eukaryotic translation initiation factor 4E confers multiallelic recessive bymovirus resistance in *Hordeum vulgare* L. *The Plant J.* 42, 912–922.
- Stoeger, E., Williams, S., Keen, D. and Christou, P. (1999) Constitutive versus seed specific expression in transgenic wheat: temporal and spatial control. *Trans. Res.* 8, 73–82.

- Sun, C., Hoglund, A. S., Olsson, H., Mangelsen, E. and Jansson, C. (2005) Antisense oligodeoxynucleotide inhibition as a potent strategy in plant biology: identification of SUSIBA2 as a transcriptional activator in plant sugar signalling. *The Plant J.* 44, 128–138.
- Sun, C., Ridderstrale, K., Hoglund, A. S., Larsson, L. G. and Jansson, C. (2007) Sweet delivery – sugar translocators as ports of entry for antisense oligodeoxynucleotides in plant cells. *The Plant J.* 52, 1192–1198.
- Suzuki, T., Eiguchi, M., Kumamaru, T., Satoh, H., Matsusaka, H., Moriguchi, K., Nagato, Y. and Kurata, N. (2008) MNU-induced mutant pools and high performance TILLING enable finding of any gene mutation in rice. *Mol. Genet. Genomics* 279, 213–223.
- Talamè, V., Bovina, R., Sanguineti, M. C., Tuberosa, R., Lundqvist, U. and Salvi, S. (2008) TILLMore, a resource for the discovery of chemically induced mutants in barley. *Plant Biotechnol. J.* 6, 477–485.
- Taylor, N. E. and Greene, E. A. (2003) PARSESNP: a tool for the analysis of nucleotide polymorphisms. *Nucleic Acids Res.* 31, 3808–3811.
- Thornecroft, D., Hosein, F., Thangavelu, M., Clark, J., Vizir, I., Burrell, M. M. and Ainsworth, C. (2003) Characterization of a gene from chromosome 1B encoding the large subunit of ADP glucose pyrophosphorylase from wheat: evolutionary divergence and differential expression of Agp2 genes between leaves and developing endosperm. *Plant Biotechnol. J.* 1, 259–270.
- Tilahun, A., Skadsen, R., Patel, M. and Kaeppler, H. (2006) The *Lem2* gene promoter of barley directs cell- and development-specific expression of *gfp* in transgenic plants. *Plant Biotechnol. J.* 4, 35–44.
- Till, B. J., Reynolds, S. H., Weil, C., Springer, N., Burtner, C., Young, K., Bowers, E., Comodo, C. A., Enns, L. C., Odden, A. R., Greene, E. A., Comai, L. and Henikoff, S. (2004) Discovery of induced point mutations in maize genes by TILLING. *BMC Plant Biol.* 4, 12.
- Tingay, S., McElroy, D., Kalla, R., Fieg, S., Wang, M., Thornton, S. and Brettell, R. (1997) *Agrobacterium tumefaciens*-mediated barley transformation. *The Plant J.* 11, 1369–1376.
- Tomita, N., Ogihara, T. and Morishita, R. (2003) Transcription factors as molecular targets: molecular mechanisms of decoy ODN and their design. *Curr. Drug Targets* 4, 603–608.
- Travella, S., Klimm, T. E. and Keller, B. (2006) RNA interference-based gene silencing as an efficient tool for functional genomics in hexaploid bread wheat. *Plant Physiol.* 142, 6–20.
- Tsutsumi, N., Kanayama, K. and Tano, S. (1992) Suppression of α -amylase gene expression by antisense oligodeoxynucleotide in barley cultured aleurone layers. *Jpn. J. Genet.* 67, 147–154.
- Vasil, V., Castillo, A. M., Fromm, M. E. and Vasil, I. K. (1992) Herbicide resistant fertile transgenic wheat plants obtained by microprojectile bombardment of regenerable embryogenic callus. *Biotechnology* 10, 667–674.
- Veena, J. H., Doerge, R. W. and Gelvin, S. (2003) Transfer of T-DNA and Vir proteins to plant cells by *Agrobacterium tumefaciens* induces expression of host genes involved in mediating transformation and suppresses host defense gene expression. *The Plant J.* 35, 219–236.
- Vickers, C. E., Xue, G. and Gresshoff, P. M. (2006) A novel *cis*-acting element, ESP, contributes to high-level endosperm-specific expression in an oat globulin promoter. *Plant Mol. Biol.* 62, 195–214.
- Wahlestedt, C., Salmi, P., Good, L., Kela, J., Johnsson, T., Hokfelt, T., Broberger, C., Porreca, F., Lai, J., Ren, K., Ossipov, M., Koshkin, A., Jakobsen, N., Skouv, J., Oerum, H., Jacobsen, M. H. and Wengel, J. (2000) Potent and nontoxic antisense oligonucleotides containing locked nucleic acids. *Proc. Natl. Acad. Sci. USA* 97, 5633–5638.
- Wan, Y. and Lemaux, P. G. (1994) Generation of large numbers of independently transformed fertile barley plants. *Plant Physiol.* 104, 37–48.
- Wang, M. B., Abbott, D. C., Upadhyaya, N. M., Jacobsen, J. V. and Waterhouse, P. M. (2001) *Agrobacterium tumefaciens*-mediated transformation of an elite Australian barley cultivar with virus resistance and reporter genes. *Aus. J. Plant Physiol.* 28, 149–156.

- Waugh, R., Leader, D. J., McCallum, N. and Caldwell, D. (2006) Harvesting the potential of induced biological diversity. *Trends Plant Sci.* 11, 71–79.
- Weeks, J. T., Anderson, O. D. and Blechl, A. E. (1993) Rapid production of multiple independent lines of fertile transgenic wheat (*Triticum aestivum*). *Plant Physiol.* 102, 1077–1084.
- Wesley, S. V., Helliwell, C. A., Smith, N. A., Wang, M. B., Rouse, D. T., Liu, Q., Gooding, P. S., Singh, S. P., Abbott, D., Stoutjesdijk, P. A., Robinson, S. P., Gleave, A. P., Green, A. G. and Waterhouse, P. M. (2001) Construct design for efficient, effective and high-throughput gene silencing in plants. *The Plant J.* 27, 581–590.
- Wiley, P. R., Tosi, P., Evrard, A., Lovegrove, A., Jones, H. D. and Shewry, P. R. (2007) Promoter analysis and immunolocalisation show that puroindoline genes are exclusively expressed in starchy endosperm cells of wheat grain. *Plant Mol. Biol.* 64, 125–136.
- Wu, H., McCormac, A. C., Elliott, M. C. and Chen, D. F. (1998) *Agrobacterium*-mediated stable transformation of cell suspension cultures of barley (*Hordeum vulgare* L.). *Plant Cell Tissue Organ Cult.* 54, 161–171.
- Wu, H., Sparks, C., Amoah, B. and Jones, H. D. (2003) Factors influencing successful *Agrobacterium*-mediated genetic transformation of wheat. *Plant Cell Rep.* 21, 659–668.
- Yan, L., Loukoianov, A., Blechl, A., Tranquilli, G., Ramakrishna, W., SanMiguel, P., Bennetzen, J. L., Echenique, V. and Dubcovsky, J. (2004) The wheat *VRN2* gene is a flowering repressor down-regulated by vernalization. *Science* 303, 1640–1644.
- Yang, B., Wen, X., Kodali, N. S., Oleykowski, C. A., Miller, C. G., Kulinski, J., Besack, D., Yeung, J. A., Kowalski, D. and Yeung, A. T. (2000) Purification, cloning, and characterization of the CEL I nuclease. *Biochemistry* 39, 3533–3541.
- Yao, Q., Cong, L., Chang, J. L., Li, K. X., Yang G. X. and He, G. Y. (2006) Low copy number gene transfer and stable expression in a commercial wheat cultivar via particle bombardment. *J. Exp. Bot.* 57, 3737–3746.
- Zhang, S., Cho, M. J., Koprek, T., Yun, R., Bregitzer, P. and Lemaux, P. G. (1999) Genetic transformation of commercial cultivars of oat (*Avena sativa* L.) and barley (*Hordeum vulgare* L.) using *in vitro* shoot meristematic cultures derived from germinated seedlings. *Plant Cell Rep.* 18, 959–966.
- Zimny, J., Becker, D., Brettschneider, R. and Lorz, H. (1995) Fertile, transgenic Triticale (X Triticosecale Wittmack). *Mol. Breed.* 1, 155–164.

Chapter 14

Genomics of Transposable Elements in the Triticeae

François Sabot and Alan H. Schulman

Abstract Triticeae genomes are structured as blocks of relatively gene-dense “islands” surrounded by long expanses of repetitive DNA. Most of the repetitive DNA is comprised of transposable elements; the greatest bulk of these are the Class I, or retrotransposons, which transpose via an RNA intermediate. The remainder is Class II DNA transposons, which move by a “cut-and-paste” mechanism. The LTR retrotransposons, which is the most abundant group of retrotransposons in the Triticeae genomes, compose 55–70 % of the genome. The precise numbers and insertion sites of members of the various families of transposons and retrotransposons in Triticeae genomes vary, which is a result of continual insertion and loss of individual copies at particular chromosomal locations. Among both the transposons and retrotransposons, non-autonomous forms are quite prevalent, among them the MITEs, SINEs, LARDs, and TRIMs. These require the proteins of autonomous forms for their mobility. Hence, the genome contains a mixture of autonomous elements, some of which contain stop codons or frameshifts inactivating translation, nonautonomous elements, and various deletion derivatives of both. The cell employs many regulatory mechanisms, including transcriptional silencing by DNA methylation and RNAi post-transcriptional silencing, to reduce transposable element propagation. Nevertheless, transposable elements have effects on Triticeae plants and their genomes over various time frames, ranging from read-through modulation of gene expression following stress activation to gene mutagenesis and growth in genome size. Genomic diversification driven by transposable element activity has made possible the exploitation of these elements as molecular markers for the Triticeae, complementing genic markers such as SNPs. With the emergence of genome sequences for members of the Triticeae and for related species such *Brachypodium distachyon*, a picture of the role of transposable elements in the evolution of genomes in the Triticeae is now emerging.

F. Sabot (✉)

MTT/BI Plant Genomics Laboratory, Institute of Biotechnology, University of Helsinki, Helsinki, Finland
e-mail: francois.sabot@gmail.com

14.1 Introduction

Triticeae genomes are known to be a kind of mosaic in which gene blocks are interspersed by repeats. Well before extensive genome sequencing, some idea of this overall pattern of organization was gleaned from a combination of cytology (Flavell et al. 1977) and DNA reassociation analysis (Ranjekar et al. 1976). Fluorescent staining of chromosomes with general reagents such as DAPI visualized frequent and fairly specific heterochromatic bands. These were shown to correspond to compacted chromatin containing an abundance of repeats (Flavell 1986). The reassociation kinetics of sheared Triticeae DNA indicated, in parallel, that about 75% of these genomes are repetitious.

The advent of DNA sequencing, combined with hybridization-based genome reconstruction calculations, made it possible to estimate the copy number of specific repetitive components. Particular retrotransposon families were shown to number in the tens of thousands in barley (Suoniemi et al. 1996) as well as in rye and wheat (Muñiz et al. 2001; Pearce et al. 1997). The ultimate cloning, sequencing, and assembly of long contiguous sequences from the Triticeae finally permitted a detailed view of how the various genome components were organized relative to one another. In a flurry of papers appearing around the turn of the millennium (Dubcovsky et al. 2001; Hudakova et al. 2001; Panstruga et al. 1998; Rostocks et al. 2002; Shirasu et al. 2000; Wicker et al. 2001; Wei et al. 2002) a pattern of gene islands of variable size surrounded by long stretches of retrotransposon insertions, which were not conserved in the homeologous chromosomes and loci of other grass species, began to emerge. The abundance of the retrotransposons as the main genome component of the Triticeae was confirmed.

These two properties, abundance and dynamic localization, are in fact a central feature of transposable elements (TEs). They are able to increase their number and change their location within the genome either by using their own enzymatic machinery or by parasitizing the machinery of other TEs. In general, they are the most abundant components of not only Triticeae, but also other large eukaryotic genomes. The coding domains of TEs form the most abundant gene families in virtually all eukaryotes. In the Triticeae, they may compose more than 80% of the genomic DNA (Fig. 14.1). As genomes become larger (Fig. 14.1), their proportion rises considerably faster than the number of genes.

The TEs are classified according their mechanism of transposition (Wicker et al. 2007). The highest level is the Class: Class I elements move using an RNA intermediate, whereas Class II transposes only via DNA (Table 14.1). The following level (Subclass) distinguishes TEs based on the number of DNA strands transferred from the original site in the genome where a particular TE is located to the acceptor site where the new insertion will appear. Below Subclass, TEs are separated in Orders, each representing a common mechanism of transposition, and then into Superfamilies, which group elements based on their structure. Family, the taxon below Superfamily, collects elements based on their sequence similarity and is the

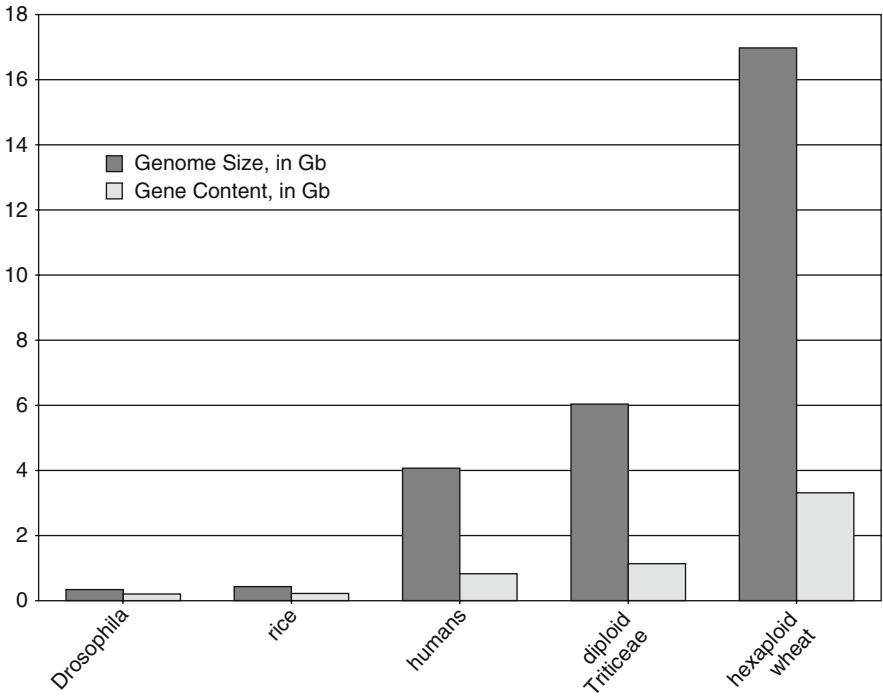


Fig. 14.1 Genome size and gene content for model species

Table 14.1 Shortened classification for the transposable elements of the Triticeae. Adapted from Wicker et al. (2007). See www.wikiposon.org for more information about the complete classification

Class	Subclass	Order	Superfamily	3-Letter
I (Retroelement)	N/A	LTR	<i>Copia</i>	RLC
			<i>Gypsy</i>	RLG
			Unclassified	RLx
		LINE	<i>RTE</i>	RIT
			<i>L1</i>	RIL
			Unclassified	RIx
		SINE	<i>tRNA</i>	RST
			<i>7SL</i>	RSL
			<i>5S</i>	RSS
			Unclassified	RSx
			Unclassified	RSx
II (DNA Transposon)	1 (double-strand DNA cut)	TIR	<i>Tc1/Mariner</i>	DTT
			<i>hAT</i>	DTA
			<i>Mutator</i>	DTM
			<i>PIF/Harbinger</i>	DTH
			<i>CACTA</i>	DTC
	2 (single-strand DNA cut)	Helitron	Unclassified	DTx
			<i>Helitron</i>	DHH
			Unclassified	DHx
			Unclassified	DHx
			Unclassified	DHx

level with which a particular element name (e.g., *BARE1* or *Bagy2*), is associated. Sometimes, another taxonomic level, the Subfamily, is used for specific groups within a family. The lowest level is that of insertion, and refers to a particular copy of an element in a specific chromosomal location.

The overall structures, domain orders, encoded enzymes, and reaction mechanisms are conserved in the various TE groups, which can be found both across the eukaryotes and in related forms in prokaryotes. The TEs therefore appear to be ancient, because the patterns of change through descent in consonance with the organismal evolutionary distance cannot be explained by horizontal transfer. At least the LINE elements, if not all the TEs, may have even appeared in LUCA (Last Universal Common Ancestor), during the RNA epoch. These LINEs have been shown to harbour an ancestral version of the reverse transcriptase domain, found elsewhere not only in LTR retrotransposons (as well as in DIRS and PLE), but also in telomerase and in the retroviruses. The DDE motifs, shared by retroviral and LTR retrotransposon integrases and by the transposases of Class II elements, form structurally-related catalytic domains with common reaction chemistries (Rice and Baker 2001; Hickman et al. 2005; Czyz et al. 2007). This provides support for the hypothesis that the presence of this domain may indicate a common origin for these groups of transposable elements (Capy et al. 1997). It should be noted that, even if eukaryotic Class II elements have related sequences in the prokaryotes, no Class I elements have been ever described in those genomes.

14.2 Structural Genomics

The TE content and its distribution between TE classes are highly variable between the various Poaceae subfamilies and their constituent tribes (see Fig. 14.2), and even sometimes between related species such as the wild relatives of rice (Piegu et al. 2006). The abundance and distribution of TEs, furthermore, differs greatly between the chromosome segments surrounding the telomeres and centromeres and the distal, gene-rich regions. Nevertheless, the overall TE share of the genome as a whole does not vary greatly within the Triticeae.

A brief structural view of the main elements encountered in the Triticeae is shown on Fig. 14.3. Within the Triticeae, the main TEs belong to the Class I (moving through an RNA intermediate). These elements (Retrotransposons, Table 14.1) are separated in 5 Orders (LTR, DIRS, PLE, LINE, and SINE). In the Triticeae, the PLE and DIRS orders have not yet been detected, so we will not consider them further. A more complete treatment of these has recently been published (Wicker et al. 2007); they are also addressed on a community website (www.wikiposon.org).

The LTR retrotransposons (Fig. 14.3A), of the Triticeae, belonging to the *Copia* and *Gypsy* superfamilies, compose between 55 and 70% of the total genomic DNA (Li et al. 2004; Paux et al. 2006; Sabot et al. 2005), depending on the genome, species, chromosome and locus considered (Fig. 14.2). They are distributed mainly in large interspersed blocks between the gene islands, where

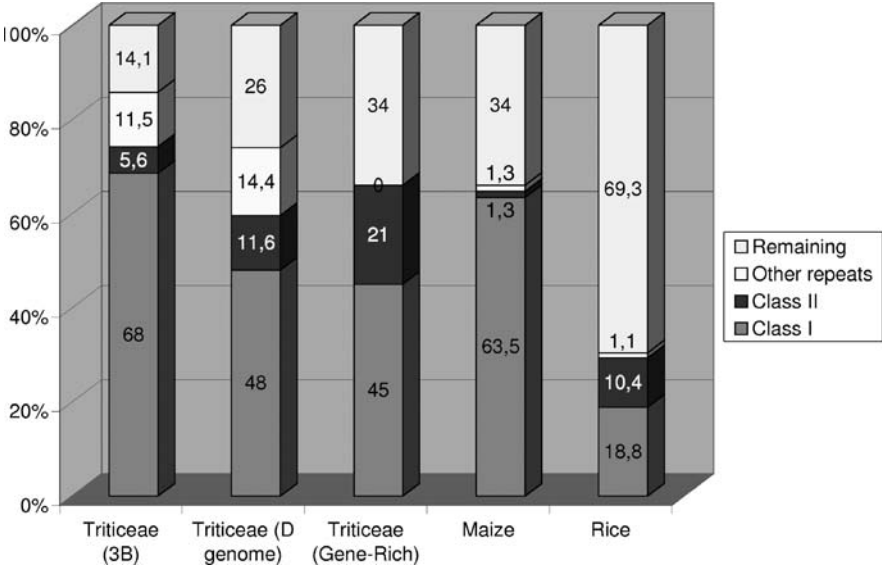


Fig. 14.2 DNA content ratio in grasses (Li et al. 2004; Paux et al. 2006; Sabot et al. 2005)

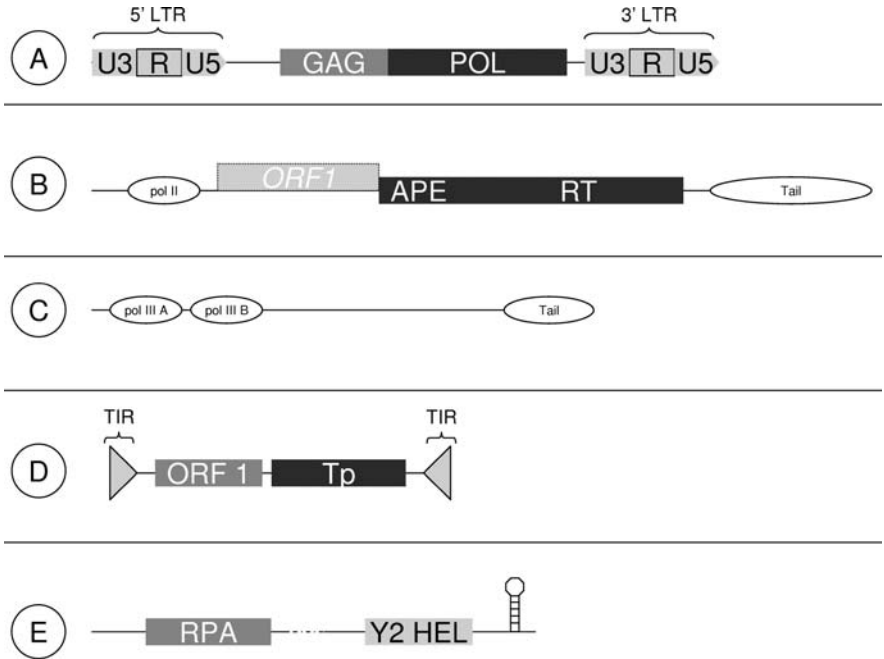


Fig. 14.3 Main TE structures. (A) LTR retrotransposons. (B) LINE elements. (C) SINE elements. (D) TIR elements (ORF1 may be missing). (E) Helitrons. See text for more explanation

they are often nested one into another. They are also present in the vicinity of genes, but more rarely. Their transposition requires the reverse transcription of their mRNA transcripts. If each transcript were to be converted to a cDNA competent for integration, the replication cycle would therefore generate as many new genomic copies as there are transcripts (see Fig. 14.4). Among the

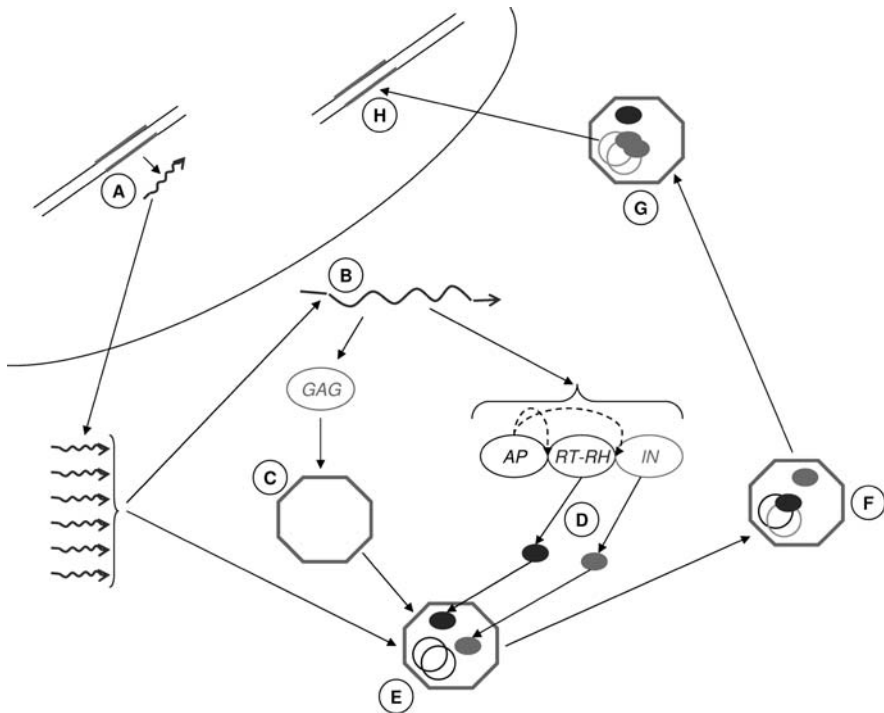


Fig. 14.4 Retrotransposition overview. (A) A genomic DNA copy is transcribed into mRNA (generally polyadenylated, but not always) by RNA polymerase II, starting at beginning of the 5' R region and stopping at the end of the 3' R region (see Fig. 14.3A). (B) This mRNA population has two fates in the cytosol: being translated into proteins and being packaged. The translation of the mRNA, containing two main ORFs (*gag* and *pol*), leads to two proteins: GAG, a structural monomer, and POL, a enzymatic polyprotein. (C) The GAG monomers polymerize in a virus-like particle (VLP), and (D) the POL is self-cleaved in RT-RNaseH and IN subunits by its own encoded AP (aspartic proteinase) subunit. (E) The mRNA (in a dimeric state), together with the RT-RNaseH and IN, is then packaged into the VLP. The specificity of the mRNA dimerization and packaging is under the control of secondary structures on the mRNA itself and of a specific motif on the GAG monomers. (F) Once packaged, the mRNA forms a loop (using the homologies of the two R regions) and is primed on the PBS (Primer Binding Site, just after the 5' LTR) using a structural RNA (generally the ^{met}tRNA) for its reverse transcription by the RT subunit. Trailing behind the RT, the RNase, which cleaves RNA hybridized to DNA, cleaves the RNA template into small fragments. Some of these serve as primers at the PPT (PolyPurineTract, located just before the 3' LTR) for the second cDNA strand synthesis. (G) and (H) Finally, the IN simultaneously generates an asymmetric double-strand break in the genome and integrates the new cDNA copy at that site. The single-stranded gap at the overhang is the repaired by the cell, leading to a "target-site duplication" (TSD) of generally 5 bp

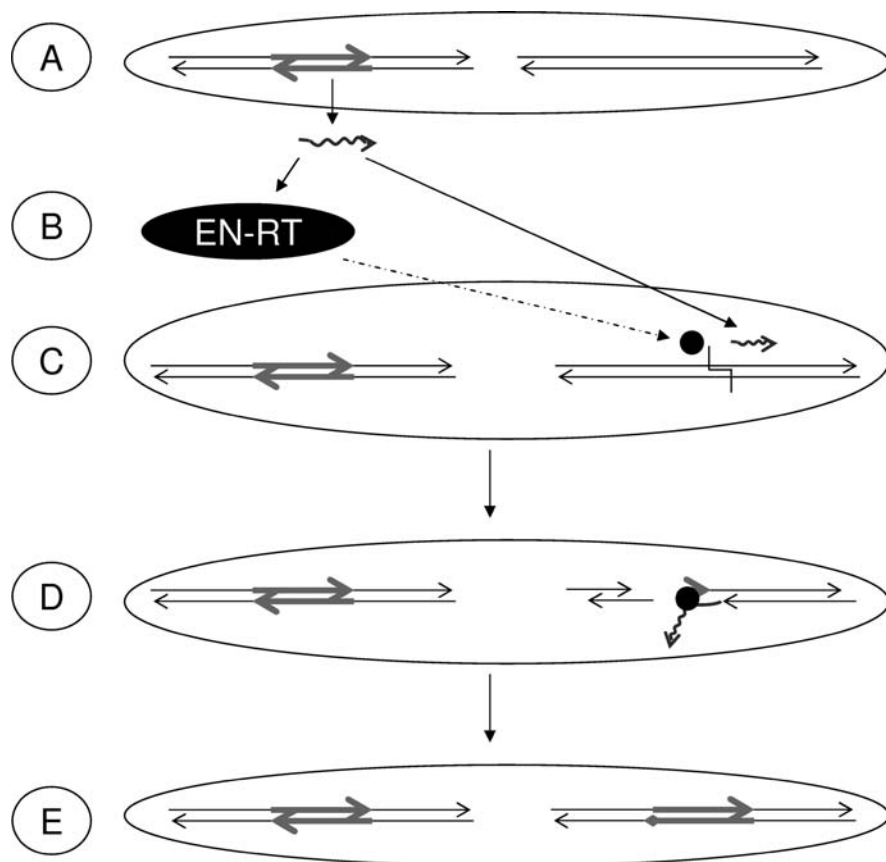


Fig. 14.5 LINE Target-Primed Reverse Transcription overview. (A) The original copy is transcribed as a standard Pol II transcript (using an internal Pol II promoter). The nucleus is shown as an oval. (B) Once translated, RT is translocated with the mRNA into the nucleus. (C) A double-strand staggered cuts at a site homologous to the "tail" motif (either performed by the endonuclease subunit or naturally occurring) is used for priming reverse transcription. (D) The reverse transcription occurs directly at the insertion site. (E) After completion of reverse transcription, a new copy is found at the new genomic location. This copy is generally 5'-truncated, the reverse transcription being stopped randomly (for the human *L1*, only 10% are complete)

Copia and *Gypsy* elements, a few families compose the main bulk in the Triticeae. These high copy number families are *BARE*, *Wis* and *Angela* (LTR retrotransposons, *Copia* for all three). The *BARE* and *Wis* elements are highly related, and the *Wis* family is sometimes refereed as the *BARE* from wheat. These three families compose from 5 to 8% of the total genomic DNA, and representatives are found everywhere in the genome.

The elements belonging to the LINE (Long Interspersed Nuclear Element) and SINE (Short Interspersed Nuclear Element) orders are present but are less studied in the Triticeae. The LINES (Fig. 14.3B) form an autonomous order,

meaning that they encode all the machinery required for their transposition (or at least some members of each family do). The main superfamilies are L1 and RTE, with the example of the *Karin* elements. Moreover, the LINES harbour an unusual internal *pol* II promoter, allowing expression of the complete LINE sequence. For these two orders, the transposition mechanism is called Target-Primed Reverse Transcription (TPRT; Fig. 14.5). The SINEs (Fig. 14.3C), in contrast, are completely non-autonomous. They derive from the abnormal reverse transcription and re-integration of RNA polymerase III products. As such, they have internal promoter boxes (Pol III A and B, Fig. 14.3C), allowing them to continue to be expressed. SINEs appear to be parasitically mobilized by LINE proteins (Garcia-Perez et al. 2007), either “stringently,” with a single obligatory partner (Dewannieux et al. 2003), or non-stringently (Fantaccione et al. 2008; Kajikawa and Okada 2002). The only element currently identified as a SINE in the Triticeae is the *Au* element, originally found in *Aegilops umbellata*.

The Class II elements are represented in Triticeae mainly by the classical DNA transposons, the TIR Order (Table 14.1; Fig. 14.3D), which move by a “copy-and-paste” mechanism (Table 14.1, Subclass 1), as shown in Fig. 14.6.

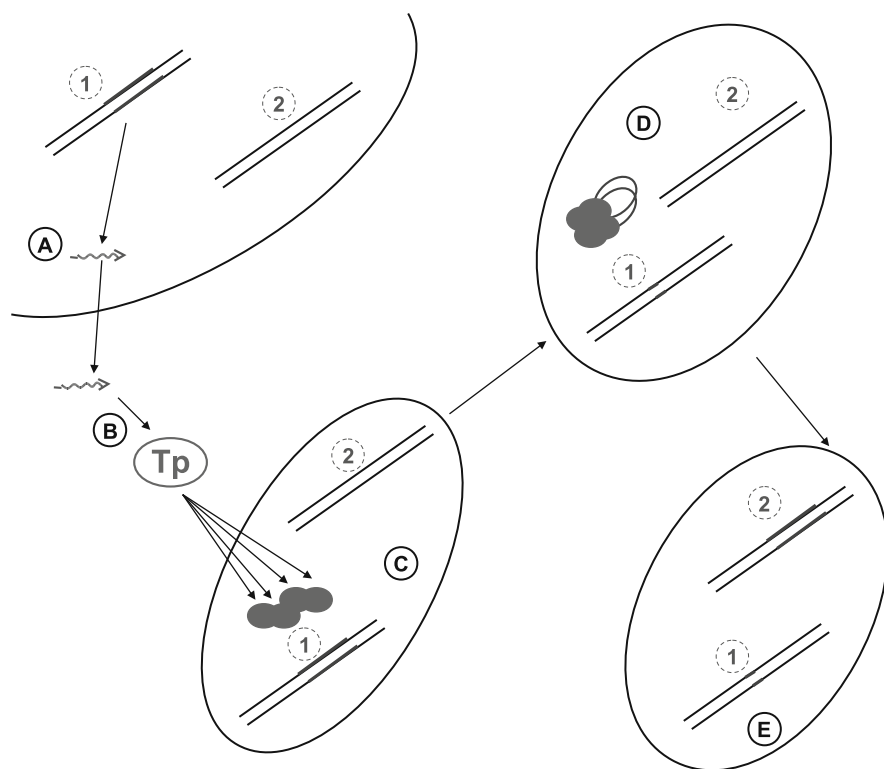


Fig. 14.6 General scheme for the TIR Order transposition by “cut-and-paste”. (A) An active copy is transcribed by cellular Pol II. (B) Once in the cytosol, this mRNA is translated into a

The main superfamilies from this Order that have been found in the Triticeae are *CACTA*, *Mutator*, and *PIF/Harbinger*. The TIR elements, especially from *CACTA*, comprise a large part of the repeats in the Triticeae, up to 20% (Wicker et al. 2003). Non-autonomous derivatives of TIR elements are widely found in the Triticeae genomes, and are referred to generally as MITEs (Miniature Inverted repeat Transposable Elements, Class II). Of the MITEs, the *Stowaway* subfamily is also highly abundant. However, the tiny size of the MITEs (around 200 bp) means that they are far better represented by their number than by their proportion of the total genome.

The *Helitron* subclass (Fig. 14.3E; Table 14.1, Subclass 2), was identified relatively recently in various organisms (Kapitonov and Jurka 2007), and has been shown in maize to be responsible for spectacular genome reorganisation (Lai et al. 2005; Morgante et al. 2005). Their mobility is thought to be driven by helicase (Kapitonov and Jurka 2007), involving a rolling-circle step in a manner analogous to the geminiviruses and bacterial *IS91* elements (Fig. 14.7). Some helicase motifs not related to cellular gene ORFs but most probably to *Helitron* elements have been detected in the currently released large sequences from Triticeae. For now, no *Crypton* (Subclass 1) or *Maverick* (Subclass 2) elements have been detected in the Triticeae, so they will not be discussed here.

14.3 Functional Genomics

14.3.1 Direct Effects

Analyses of EST data have shown that a minimum of 1% of expressed RNA comes from TEs. The TE sequences of EST databases have led researchers to overestimate gene numbers, because only genes were expected to be expressed at such high levels. One can estimate that about 134,000 mRNAs are present per Triticeae cell. We have calculated this based on observations that total RNA yields are about equal to DNA yields and 1% by weight of total RNA is poly(A) RNA, on a DNA content in barley of 10.787 pg per cell, and on the average gene size of rice, which is 1,514 nt when only the exons are considered (<http://www.tigr.org/tdb/e2k1/osal/riceInfo/info.shtml>). This estimate is consistent with observations of 40–80 thousand transcripts per ng of total RNA from rice

Fig.14.6 (continued) transposase (Tp). (C) Transposases are then relocated to the nucleus, where they will link to the TIRs of inserted transposon copies. (D) The transposases will perform a double-stranded cut (alone or with a host protein performing the second cut), releasing the double-strand DNA copy, and leaving excision footprints. (E) The excised double-stranded DNA is then re-inserted, at a fairly nearby genomic location, at a staggered double-stranded break made by the transposase. Once repaired by host enzymes, the break generates TSDs

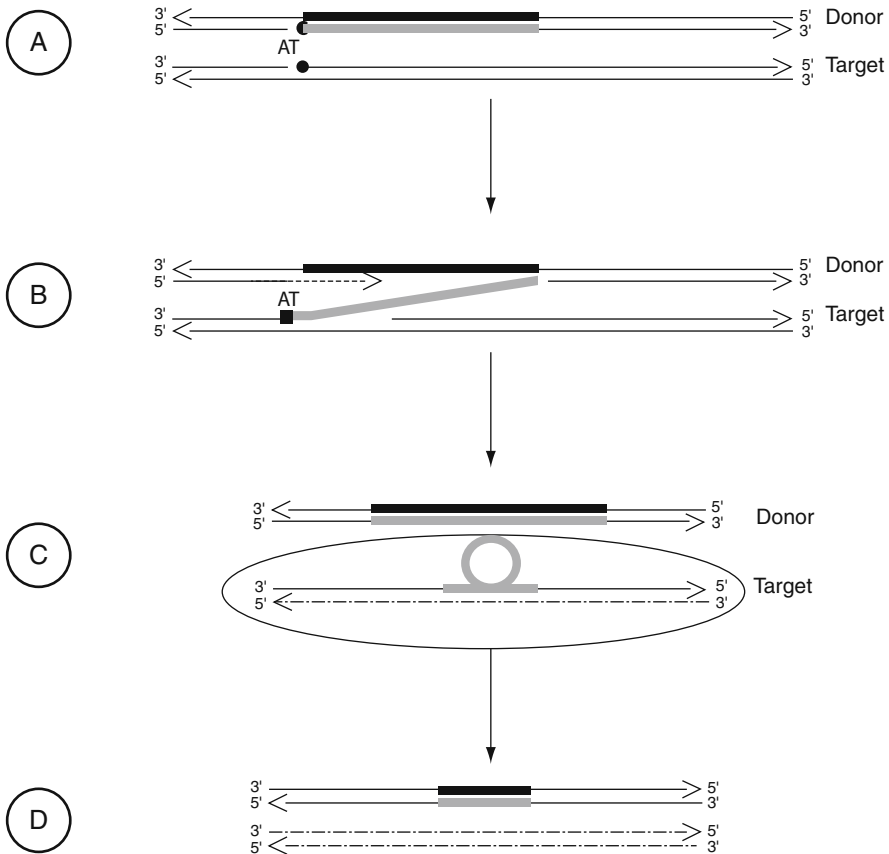


Fig. 14.7 A putative "Rolling-Circle" transposition mechanism for Helitrons (based on the bacterial element IS91). (A) The element (in *dark grey*) may be autonomous or non-autonomous. Two proteins similar to replication protein A (*black dots*; see Fig. 14.3E, RPA) cut the donor and acceptor sites, and (B) make a ligation between 3' donor strand and 5' acceptor strand. Replication starts at free 3' end of the donor site, and allows the helitron's single-strand displacement. When the palindromic structure and the 3' sequence are recognized, a cut is made after the CTRR and (C) the strand is transferred. (D) The heteroduplex is resolved during the following DNA replication round

endospore being present for highly expressed genes (Ohdan et al. 2005), and of 15,000 genes per tissue being expressed in barley (Druka et al. 2006).

Given 100,000 transcripts per cell and 1% from TEs, there are about 1,000 TEs transcripts and roughly 800 from retrotransposons. For each pair of transcripts packaged into a VLP, a retrotransposon is theoretically able to create a new insertion. If no post-transcriptional controls were operating and transcriptional expression would be correlated to the insertion of new copies TEs, we would expect to see 400 new insertions per cell cycle and an increase in genome size of 400 kb (given a mean size of 10 kb per element). Even more may

be expected, because we have considered the mRNA pool size and not the total number of transcripts generated per cell cycle. Even so, if the average cell is 100 μm long and the average Triticeae plant 1 m tall, one would expect a minimum of 10,000 cell divisions seed to seed and a consequent 4×10^9 bp increase in genome size per plant generation, roughly equivalent to a doubling of the basic genome size (5×10^9 bp in barley). Such an explosive growth, let alone the potential gene mutations linked to those new insertions, is not seen. This rough calculation highlights the dissociation between the expression and the transposition of TEs, and emphasizes the many levels of control.

The first level of transpositional attenuation can be found within the TE sequences themselves. On this level, transcripts are produced that cannot in and of themselves drive the life cycle through to a new insertion. Most of the known TE insertions within Triticeae genomes are either non-functional copies of autonomous elements, containing ORFs that are peppered with stop codons and indels, or non-autonomous elements. The non-autonomous elements are those partially or even completely lacking coding capacity (as defined in Sabot and Schulman 2006). Non-autonomous groups of elements, such as MITEs, LARDs, and TRIMs (Kalendar et al. 2004; Kikuchi et al. 2003; Witte et al. 2001), must parasitize their autonomous partners, another TE group, to ensure their mobility. Some non-autonomous groups are highly efficient at parasitism; it appears, for example, that the most numerous retroelements in barley are members of the *BARE2* subfamily, which cannot produce the GAG essential for packaging (Tanskanen et al. 2007).

A more extreme example of loss of coding capacity are the *Sukkula* elements (Kalendar et al. 2004), which are LARDs. These LTR retrotransposons are non-autonomous giants (14 kb, no ORFs within), but are the most expressed TEs in barley, with almost 10% of the total TE mRNAs derived from them (Sabot et al. unpublished data). Despite their active expression as RNA, few *Sukkula* transcripts, probably less than 0.01% of the total, are ultimately integrated into the genome as new copies. Indeed, based on their Activity Index (Sabot et al. 2005), which allows comparison of TE groups based on insertion patterns, they are among the least active retrotransposons in the Triticeae.

Although TE transcription may appear to be extensive, there is nonetheless evidence for transcriptional silencing of these genetic components. This represents a second point of control. Many TEs are located in heterochromatin, and evidence has accumulated for the targeting of TEs by mechanisms in directing DNA methylation, histone modification, and heterochromatization (Bäurle et al. 2007; Cheng et al. 2006; Lippman et al. 2003; Penterman et al. 2007; Tran et al. 2005; Woodhouse et al. 2006). As a secondary effect of the methylation of TEs, they can accumulate additional mutations, in particular C to T transversions.

A third level of control over TEs includes host mechanisms that work post-transcriptionally to block expression, such as RNAi. Various TE transcripts may be recognized by post transcriptional regulatory mechanisms, such as that

involving the Dicer RNase and microRNAs (Baulcombe 2006; Jones-Rhoades et al. 2006), due to either their containing long double-stranded regions or their being expressed at exceptionally high levels. Long double-stranded regions can result from inverted, nested insertions, whereas antisense transcripts can result from transcriptional read-through by cellular promoters. Evidence has accumulated for the post-transcriptional silencing of both Class I and Class II transposable elements in a variety of plants (Neumann et al. 2007; Ronemus et al. 2006; Rudenko et al. 2003; Zhang et al. 2006).

An additional level of control may be acting on autonomous TEs, akin to interference seen in the viral world (García-Arriaza et al. 2006; Li and Pattnaik 1997; Qiu et al. 2001; Roux et al. 1991). Non-autonomous TEs may interfere with the propagation of autonomous partners by means that include being more effectively packaged by their proteins, by interfering with protein multimer formation, or by setting off RNAi pathways. Parasitism by non-autonomous TEs has the potential to drive the autonomous partners to extinction.

14.3.2 Effects on Genes, Sequence Chimeras, and Gene Regulation

Outside of their effects on their own expression and on the size of the genome, TEs may also have effects in *cis*- and *trans*- on the expression of cellular genes. About half of the known MITE insertions can be found within the regulatory regions of the genes (1 kb upstream or downstream) or within introns (Sabot et al. 2005). Such an association may have two origins. First, MITEs may insert preferentially in relaxed chromatin regions, near expressed genes. Alternatively, those MITE insertions that are near to genes may be selectively maintained by their potential regulatory effects. Because they are able to form stable secondary structures, a role for MITEs in the formation of chromatin boundaries or MARs has been proposed (Anderson et al. 2002), as earlier seen for *Gypsy* elements in *Drosophila*. (Nabirochkin et al. 1998) and possibly for retrotransposons in *Petunia* (Galliano et al. 1995).

Some ESTs appear to derive from chimeras between cellular genes and TE transcripts. Such chimeras may appear by read-through transcription from either the cellular gene or the TE. Alternatively, they may result from insertions of TEs into an intron, blocking correct splicing, or into 5' or 3' untranslated regions. Read-through from TE transcription may be seen only through stress or "genomic shock" activation of the TE promoters (Kashkush et al. 2003). Activating conditions include chemical treatments, *in vitro* culture, and interspecific hybridization. An important potential consequence, seen in the Triticeae, is the suppression of the cellular gene because of the read-through generation of antisense RNAs to the gene (Kashkush et al. 2003). Likewise, there is evidence for the role of interfering RNA in silencing retrotransposons (Huettel et al. 2007; Neumann et al. 2007).

14.4 Comparative Genomics

The availability of long contiguous sequences from orthologous regions of related genomes has made comparative genomics possible, as described in the Introduction. Information can be gleaned, not only on the expansion of genomes and the evolution of particular cellular genes, but also on the TEs themselves. Even if the species within the Triticeae are sufficiently close to allow hybrids in some cases (such as triticale) and to conserve considerable genetic colinearity and microsynteny, the TE history of each genome is unique, even in highly related species (Chantret et al. 2005; Sabot et al. 2005; Tanskanen et al. 2007). Recombination driven by TEs may create chromosomal incompatibilities, precipitating speciation.

The overall composition of the TE complement in the various Triticeae genomes, as measured by the relative abundance of the different Orders, is generally quite well conserved (Fig. 14.2). However, the particular families found in high copy number differ from species to species. We have shown that the main difference between the cultivated barley *Hordeum vulgare*, together with its wild ancestor *H. spontaneum*, compared to other barley species is the abundance of the *BARE* family and especially of the *BARE2* subfamily (Tanskanen et al. 2007; Vicient et al. 1999). Even within *H. spontaneum*, sharp dichotomies in *BARE* copy number are seen across ecological boundaries (Kalendar et al. 2000).

Clearly, TE “bursts” have occurred within *Hordeum*, though in the absence of paired LTRs from many species it is difficult to reconstruct the timing of these events (Kronmiller and Wise 2007; SanMiguel et al. 1998). Similar bursts of TE integration are apparent in the Triticeae A genome, which is probably larger than the B and D genomes as a result (Sabot et al. 2005), whereas the *Morgane* family in particular amplified in the D genome of wheats (Sabot et al. 2006). The mean size of the Triticeae genomes is nevertheless conserved, perhaps because of the various mechanisms controlling TE activity as well as several means of TE loss through recombination (Rabinowicz 2000; Shirasu et al. 2000; Vicient et al. 1999; Vitte and Panaud 2003).

14.5 Exploitation as Molecular Markers

Several features of transposable elements have made them ideally suited for development as molecular markers (Kalendar and Schulman 2006; Schulman 2007). An advantage of the retrotransposons in particular is that they do not excise to transpose, and leave the original insertion intact. The MITEs as well are rarely seen to excise. One possible explanation is that following excision, a gap repair mechanism using homologous recombination may, due to the short size of the MITE, be able to repair and re-create the excised MITE, using the MITE site on the corresponding chromosome. Therefore, unlike marker

systems based on single nucleotide changes (RFLP, AFLP, SNPs) and those based on changes in the number of simple sequence repeats or microsatellites, TE markers do not suffer from homoplasy. This means that the derived state cannot be confused with an ancestral state. Furthermore, even LTR-LTR recombination leaves the LTR-genome joint intact, making Class I TEs very stable as markers. They create large changes in the genome that can easily be cloned and are amenable to PCR amplification due to the conserved sequences in the TEs. The TEs are highly abundant and distributed throughout the genome. Moreover, the separate evolutionary histories and windows of activity for different TE families means that a TE marker system can be chosen to match the desired phylogenetic resolution.

Over the last decade, several different retrotransposon-based marker systems have been introduced. The S-SAP or SSAP (Sequence-Specific Amplified Polymorphism) method produces PCR products between retrotransposons and adapters ligated at restriction sites (Waugh et al. 1997). Essentially the same technique was named TD (Transposon Display) when applied to transposons rather than retrotransposons a year later (Van den Broeck et al. 1998). The IRAP method amplifies products between two retrotransposons, whereas REMAP amplifies between a retrotransposon and a simple sequence repeat or microsatellite (Kalendar et al. 1999). All three methods are dominant, meaning that the allelic state corresponding to the absence of a TE insertion at a particular locus cannot be detected positively. When the flanks of a TE insertion are identified, the RBIP (retrotransposon-based insertional polymorphism) method (Flavell et al. 1998) or its chip-based implementation (Flavell et al. 2003) can be used. A variant of the RBIP, the Insertion Site Based Polymorphism (ISBP) method, has been developed recently in wheat (Paux et al. 2006) and has been applied to evolutionary studies and breeding programs (Santenac et al. 2009).

Our recent survey of PubMed revealed that, since their introduction in 1997, over 190 publications have appeared in which Class I or Class II TEs served as molecular markers. Laboratory guides have been published for their practice (Casa et al. 2004; Kalendar and Schulman 2006; Schulman et al. 2004). In the Triticeae, the methods have been applied to barley (Kalendar et al. 1999, 2000, 2004; Leigh et al. 2003; Manninen et al. 2000; Soleimani et al. 2005; Waugh et al. 1997), wheat and its relatives (Boyko et al. 2002; Nagy et al. 2006; Queen et al. 2004), rye (Nagy and Leleý 2003). Although most of the insertions serving as markers are not themselves responsible for the trait they track, there are some well-known exceptions. One is the wheat null Glu-1A allele for high molecular weight glutenin, in which a *Wis2* element is inserted (Harberd et al. 1987). Another, though not in the Triticeae, nonetheless is interesting historically: *Gypsy* element insertions into *Vitis* transcription factors have led to white grapes (Kobayashi et al. 2004; This et al. 2007). We fully expect that as more genome sequence data emerge, TE markers will play a major role not only in breeding, biodiversity, and germplasm resources applications, but also in genomics efforts to link physical with recombinational maps.

14.6 Conclusions

TEs are the main actors in Triticeae genomes, both today and in the past. They reshape the genome and affect plant evolution by various processes: genome growth through integration; recombinational loss and rearrangements; insertional inactivation of genes; triggering of chromatin methylation silencing nearby genes; read-through activation and silencing of nearby genes; donation of promoter functions to cellular genes. The presence of both autonomous and non-autonomous groups may set up a complex dynamic of parasitism and interference that may be under selective pressure. The disjunction between TE transcriptional rates and insertional frequencies points to multiple levels of cellular regulation acting upon them. The ability of TE promoters to be activated by stresses connects TEs to general cellular signalling pathways. Although many genomic copies in the Triticeae and elsewhere are non-autonomous or even fully inactive fossils, TEs are not inert junk; they act within and upon the genome and they act upon the each other. Hence, the Triticeae and other genomes are best seen as ecosystems with predators and parasites. The study of transposable elements will provide insights and keys to the mechanisms driving Triticeae evolution, human selection, and behaviour.

References

- Anderson, O.D., Larka, L., Christoffers, M.J., McCue, K.F. and Gustafson, J.P. (2002) Comparison of orthologous and paralogous DNA flanking the wheat high molecular weight glutenin genes: sequence conservation and divergence, transposon distribution, and matrix-attachment regions. *Genome* 45, 367–380.
- Baulcombe, D.C. (2006) Short silencing RNA: the dark matter of genetics? *Cold Spring Harbor Symp. Quant. Biol.* 71, 13–20.
- Bäurle, I., Smith, L., Baulcombe, D.C. and Dean, C. (2007) Widespread role for the flowering-time regulators FCA and FPA in RNA-mediated chromatin silencing. *Science* 318, 109–112.
- Boyko, E., Kalendar, R., Korzun, V., Gill, B. and Schulman, A.H. (2002) Combined mapping of *Aegilops tauschii* by retrotransposon, microsatellite, and gene markers. *Plant Mol. Biol.* 48, 767–790.
- Capy, P., Langin, T., Higuete, D., Maurer, P. and Bazin, C. (1997) Do the integrases of LTR-retrotransposons and class II element transposases have a common ancestor? *Genetica* 100, 63–72.
- Casa, A.M., Nagel, A. and Wessler, S.R. (2004) MITE display. *Methods Mol. Biol.* 260, 175–188.
- Chantret, N., Salse, J., Sabot, F., Rahman, S., Bellec, A., Laubin, B., Dubois, I., Dossat, C., Sourdille, P., Joudrier, P., Gautier, M.F., Cattolico, L., Beckert, M., Aubourg, S., Weissenbach, J., Caboche, M., Bernard, M., Leroy, P. and Chalhou, B. (2005) Molecular basis of evolutionary events that shaped the hardness locus in diploid and polyploid wheat species (*Triticum* and *Aegilops*). *Plant Cell* 17, 1033–1045.
- Cheng, C., Daigen, M. and Hirochika, H. (2006) Epigenetic regulation of the rice retrotransposon Tos17. *Mol. Genet. Genomics* 276, 378–390.
- Czyz, A., Stillmock, K.A., Hazuda, D.J. and Reznikoff, W.S. (2007) Dissecting Tn5 transposition using HIV-1 integrase diketoacid inhibitors. *Biochemistry* 46, 10776–10789.

- Dewannieux, M., Esnault, C. and Heidmann, T. (2003) LINE-mediated retrotransposition of marked Alu sequences. *Nat. Genet.* 35, 41–48.
- Druka, A., Muehlbauer, G., Druka, I., Caldo, R., Baumann, U., Rostoks, R., Schreiber, A., Wise, R., Close, T., Kleinhofs, A., Graner, A., Schulman, A., Langridge, P., Sato, K., Hayes, P., McNicol, J., Marshall, D. and Waugh, R. (2006) An atlas of gene expression from seed to seed through barley development. *Funct. Integr. Genomics* 6, 202–211.
- Dubcovsky, J., Ramakrishna, W., SanMiguel, P.J., Busso, C.S., Yan, L., Shiloff, B.A. and Bennetzen, J.L. (2001) Comparative sequence analysis of colinear barley and rice bacterial artificial chromosomes. *Plant Physiol.* 125, 1342–1353.
- Fantaccione, S., Woodrow, P. and Pontecorvo, G. (2008) Identification of a family of SINES and LINEs in the *Pipistrellus kuhli* genome: a new structural and functional symbiotic relationship. *Genomics* 91, 178–185.
- Flavell, A.J., Knox, M.R., Pearce, S.R. and Ellis, T.H.N. (1998) Retrotransposon-based insertion polymorphisms (RBIP) for high throughput marker analysis. *Plant J.* 16, 643–650.
- Flavell, A.J., Bolshakov, V.N., Booth, A., Jing, R., Russell, J., Ellis, T.H. and Isaac, P. (2003) A microarray-based high throughput molecular marker genotyping method: the tagged microarray marker (TAM) approach. *Nucl. Acids Res.* 31, e115.
- Flavell, R.B., Rimpau, J. and Smith, D.B. (1977) Repeated sequence DNA relationships in four cereal genomes. *Chromosoma* 63, 205–222.
- Flavell, R.B. (1986) Repetitive DNA and chromosome evolution in plants. *Phil. Trans. R. Soc. Lond. B* 312, 227–242.
- Galliano, H., Muller, A.E., Lucht, J.M. and Meyer, P. (1995) The transformation booster sequence from *Petunia hybrida* is a retrotransposon derivative that binds to the nuclear scaffold. *Mol. Gen. Genet.* 247, 614–622.
- García-Arriaza, J., Ojosnegros, S., Dávila, M., Domingo, E. and Escarmís, C. (2006) Dynamics of mutation and recombination in a replicating population of complementing, defective viral genomes. *J. Mol. Biol.* 360, 558–572.
- García-Perez, J.L., Doucet, A.J., Bucheton, A., Moran, J.V. and Gilbert, N. (2007) Distinct mechanisms for trans-mediated mobilization of cellular RNAs by the LINE-1 reverse transcriptase. *Genome Res.* 17, 602–611.
- Harberd, N.P., Flavell, R.B. and Thompson, R.D. (1987) Identification of a transposon-like insertion in a Glu-1 allele of wheat. *Mol. Gen. Genet.* 209, 326–332.
- Hickman, A., Perez, Z., Zhou, L., Musungarimi, P., Ghirlando, R., Hinshaw, J., Craig, N. and Dyda, F. (2005) Molecular architecture of a eukaryotic DNA transposase. *Nat. Struct. Biol.* 12, 715–721.
- Hudakova, S., Michalek, W., Presting, G.G., ten Hoopen, R., dos Santos, K., Jasencakova, Z. and Schubert, I. (2001) Sequence organization of barley centromeres. *Nucl. Acids Res.* 29, 5029–5035.
- Huettel, B., Kanno, T., Daxinger, L., Bucher, E., van der Winden, J., Matzke, A.J. and Matzke, M. (2007) RNA-directed DNA methylation mediated by DRD1 and Pol IVb: a versatile pathway for transcriptional gene silencing in plants. *Biochim. Biophys. Acta* 1769, 358–374.
- Jones-Rhoades, M.W., Bartel, D.P. and Bartel, B. (2006) MicroRNAs and their regulatory roles in plants. *Ann. Rev. Plant Physiol.* 57, 19–53.
- Kajikawa, M. and Okada, N. (2002) LINEs mobilize SINES in the eel through a shared 3' sequence. *Cell* 111, 433–444.
- Kalendar, R., Grob, T., Regina, M., Suoniemi, A. and Schulman, A.H. (1999) IRAP and REMAP: two new retrotransposon-based DNA fingerprinting techniques. *Theor. Appl. Genet.* 98, 704–711.
- Kalendar, R., Tanskanen, J., Immonen, S., Nevo, E. and Schulman, A.H. (2000) Genome evolution of wild barley (*Hordeum spontaneum*) by *BARE-1* retrotransposon dynamics in response to sharp microclimatic divergence. *Proc. Natl. Acad. Sci. USA* 97, 6603–6607.

- Kalendar, R., Vicient, C.M., Peleg, O., Ananthawat-Jonsson, K., Bolshoy, A. and Schulman, A.H. (2004) LARD retroelements: novel, non-autonomous components of barley and related genomes. *Genetics* 166, 1437–1450.
- Kalendar, R. and Schulman, A. (2006) IRAP and REMAP for retrotransposon-based genotyping and fingerprinting. *Nat. Protoc.* 1, 2478–2484.
- Kapitonov, V.V. and Jurka, J. (2007) Helitrons on a roll: eukaryotic rolling-circle transposons. *Trends Genet.* 23, 521–529.
- Kashkush, K., Feldman, M. and Levy, A.A. (2003) Transcriptional activation of retrotransposons alters the expression of adjacent genes in wheat. *Nat. Genet.* 32, 102–106.
- Kikuchi, K., Terauchi, K., Wada, M. and Hirano, H.Y. (2003) The plant MITE mPing is mobilized in anther culture. *Nature* 421, 167–170.
- Kobayashi, S., Goto-Yamamoto, N. and Hirochika, H. (2004) Retrotransposon-induced mutations in grape skin color. *Science* 304, 982.
- Kronmiller, B.A. and Wise, R.P. (2007) TE nest: automated chronological annotation and visualization of nested plant transposable elements. *Plant Physiol.* 146, 45–59.
- Lai, J., Li, Y., Messing, J. and Dooner, H.K. (2005) Gene movement by Helitron transposons contributes to the haplotype variability of maize. *Proc. Natl. Acad. Sci. USA* 102, 9068–9073.
- Leigh, F., Kalendar, R., Lea, V., Lee, D., Donini, P. and Schulman, A.H. (2003) Comparison of the utility of barley retrotransposon families for genetic analysis by molecular marker techniques. *Mol. Genet. Genomics* 269, 464–474.
- Li, T. and Pattnaik, A.K. (1997) Replication signals in the genome of vesicular stomatitis virus and its defective interfering particles: identification of a sequence element that enhances DI RNA replication. *Virology* 232, 248–259.
- Li, W., Zhang, P., Fellers, J.P., Friebe, B. and Gill, B.S. (2004) Sequence composition, organization, and evolution of the core Triticeae genome. *Plant J.* 40, 500–511.
- Lippman, Z., May, B., Yordan, C., Singer, T. and Martienssen, R. (2003) Distinct mechanisms determine transposon inheritance and methylation via small interfering RNA and histone modification. *PLoS Biol.* 1, E67.
- Manninen, O., Kalendar, R., Robinson, J. and Schulman, A.H. (2000) Application of *BARE-1* retrotransposon markers to map a major resistance gene for net blotch in barley. *Mol. Gen. Genet.* 264, 325–334.
- Morgante, M., Brunner, S., Pea, G., Fengler, K., Zuccolo, A. and Rafalski, A. (2005) Gene duplication and exon shuffling by helitron-like transposons generate intraspecies diversity in maize. *Nat. Genet.* 37, 997–1002.
- Muñiz, L.M., Cuadrado, A., Jouve, N. and Gonzales, J.M. (2001) The detection, cloning, and characterisation of WIS 2-1A retrotransposon-like sequences in *Triticum aestivum* and *xTriticosecale* Wittmack and an examination of their evolution in related Triticeae. *Genome* 44, 978–989.
- Nabirochkin, S., Ossokinam, M. and Heidmann, T. (1998) A nuclear matrix/scaffold attachment region co-localizes with the gypsy retrotransposon insulator sequence. *J. Biol. Chem.* 273, 2473–2479.
- Nagy, E.D. and Lelley, T. (2003) Genetic and physical mapping of sequence-specific amplified polymorphic (SSAP) markers on the IRS chromosome arm of rye in a wheat background. *Theor. Appl. Genet.* 107, 1271–1277.
- Nagy, E.D., Molnar, I., Schneider, A., Kovacs, G. and Molnar-Lang, M. (2006) Characterization of chromosome-specific S-SAP markers and their use in studying genetic diversity in *Aegilops* species. *Genome* 49, 289–296.
- Neumann, P., Yan, H. and Jiang, J. (2007) The centromeric retrotransposons of rice are transcribed and differentially processed by RNA interference. *Genetics* 176, 749–761.
- Ohdan, T., Francisco, P.B., Jr., Sawada, T., Hirose, T., Terao, T., Satoh, H. and Nakamura, Y. (2005) Expression profiling of genes involved in starch synthesis in sink and source organs of rice. *J. Exp. Bot.* 56, 3229–3244.

- Panstruga, R., Büschges, R., Piffanelli, P. and Schulze-Lefert, P. (1998) A contiguous 60 kb genomic stretch from barley reveals molecular evidence for gene islands in a monocot genome. *Nucl. Acids Res.* 26, 1056–1062.
- Paux, E., Roger, D., Badaeva, E., Gay, G., Bernard, M., Sourdille, P. and Feuillet, C. (2006) Characterizing the composition and evolution of homoeologous genomes in hexaploid wheat through BAC-end sequencing on chromosome 3B. *Plant J.* 48, 463–474.
- Pearce, S.R., Harrison, G., Heslop-Harrison, J.S., Flavell, A.J. and Kumar, A. (1997) Characterization and genomic organization of Ty1-*copia* group retrotransposons in rye (*Secale cereale*). *Genome* 40, 1–9.
- Penterman, J., Uzawa, R. and Fischer, R.L. (2007) Genetic Interactions between DNA demethylation and methylation in *Arabidopsis thaliana*. *Plant Physiol.* 145, 1549–1557.
- Piegu, B., Guyot, R., Picault, N., Roulin, A., Saniyal, A., Kim, H.I., Collura, K., Brar, D.S., Jackson, S.A., Wing, R.A. and Panaud, O. (2006) Doubling genome size without polyploidization: dynamics of retrotransposition-driven genomic expansions in *Oryza australiensis*, a wild relative of rice. *Genome Res.* 16, 1262–1269.
- Qiu, W., Park, J.W., Jackson, A.O. and Scholthof, H.B. (2001) Retention of a small replicase gene segment in tomato bushy stunt virus defective RNAs inhibits their helper-mediated trans-accumulation. *Virology* 281, 51–60.
- Queen, R.A., Gribbon, B.M., James, C., Jack, P. and Flavell, A.J. (2004) Retrotransposon-based molecular markers for linkage and genetic diversity analysis in wheat. *Mol. Gen. Genomics* 271, 91–97.
- Rabinowicz, P.D. (2000) Are obese plant genomes on a diet? *Genome Res.* 10, 893–894.
- Ranjekar, P.K., Pallotta, D. and Lafontaine, J.G. (1976) Analysis of the genome of plants. II. Characterization of repetitive DNA in barley (*Hordeum vulgare*) and wheat (*Triticum aestivum*). *Biochim. Biophys. Acta* 425, 30–40.
- Rice, P.A. and Baker, T.A. (2001) Comparative architecture of transposase and integrase complexes. *Nat. Struct. Biol.* 8, 302–307.
- Ronemus, M., Vaughn, M.W. and Martienssen, R.A. (2006) MicroRNA-targeted and small interfering RNA-mediated mRNA degradation is regulated by argonaute, dicer, and RNA-dependent RNA polymerase in *Arabidopsis*. *Plant Cell* 18, 1559–1574.
- Rostocks, N., Park, Y.J., Ramakrishna, W., Ma, J., Druka, A., Shiloff, B.A., SanMiguel, P.J., Jiang, Z., Brueggeman, R., Sandhu, D., Gill, K., Bennetzen, J.L. and Kleinhofs, A. (2002) Genomic sequencing reveals gene content, genomic organization, and recombination relationships in barley. *Funct. Integr. Genomics* 2, 51–59.
- Roux, L., Simon, A.E. and Holland, J.J. (1991) Effects of defective interfering viruses on virus replication and pathogenesis *in vitro* and *in vivo*. *Adv. Virus Res.* 40, 181–211.
- Rudenko, G.N., Ono, A. and Walbot, V. (2003) Initiation of silencing of maize MuDR/Mu transposable elements. *Plant J.* 33, 1013–1025.
- Sabot, F., Guyot, R., Wicker, T., Chantret, N., Laubin, B., Chalhoub, B., Leroy, P., Sourdille, P. and Bernard, M. (2005) Updating of transposable element annotations from large wheat genomic sequences reveals diverse activities and gene associations. *Mol. Gen. Genomics* 274, 119–130.
- Sabot, F. and Schulman, A.H. (2006) Parasitism and the retrotransposon life cycle in plants: a hitchhiker's guide to the genome. *Heredity* 97, 381–388.
- Sabot, F., Sourdille, P., Chantret, N. and Bernard, M. (2006) *Morgane*, a new LTR retrotransposon group, and its subfamilies in wheats. *Genetica* 128, 439–447.
- Saintenac, C., Falque, M., Martin, O.C., Paux, E., Feuillet, C. and Sourdille, P. (2009) Detailed recombination studies along chromosome 3B provide new insights on crossover distribution in wheat (*Triticum aestivum* L.). *Genetics* 181, 393–403.
- SanMiguel, P., Gaut, B.S., Tikhoniv, A., Nakajima, Y. and Bennetzen, J.L. (1998) The paleontology of intergene retrotransposons in maize. *Nat. Genet.* 20, 43–45.
- Schulman, A.H., Flavell, A.J. and Ellis, T.H.N. (2004) The application of LTR retrotransposons as molecular markers in plants. *Methods Mol. Biol.* 260, 145–173.

- Schulman, A.H. (2007) Molecular markers to assess genetic diversity. *Euphytica* 158, 313–321.
- Shirasu, K., Schulman, A.H., Lahaye, T. and Schulze-Lefert, P. (2000) A contiguous 66 kb barley DNA sequence provides evidence for reversible genome expansion. *Genome Res.* 10, 908–915.
- Soleimani, V.D., Baum, B.R. and Johnson, D.A. (2005) Genetic diversity among barley cultivars assessed by sequence-specific amplification polymorphism. *Theor. Appl. Genet.* 110, 1290–1300.
- Suoniemi, A., Ananthawat-Jónsson, K., Arna, T. and Schulman, A.H. (1996) Retrotransposon *BARE-1* is a major, dispersed component of the barley (*Hordeum vulgare* L.) genome. *Plant Mol. Biol.* 30, 1321–1329.
- Tanskanen, J.A., Sabot, F., Vicient, C. and Schulman, A.H. (2007) Life without GAG: the *BARE-2* retrotransposon as a parasite's parasite. *Gene* 390, 166–174.
- This, P., Lacombe, T., Cadle-Davidson, M. and Owens, C.L. (2007) Wine grape (*Vitis vinifera* L.) color associates with allelic variation in the domestication gene *VvmybA1*. *Theor. Appl. Genet.* 114, 723–730.
- Tran, R.K., Zilberman, D., de Bustos, C., Ditt, R.F., Henikoff, J.G., Lindroth, A.M., Delrow, J., Boyle, T., Kwong, S., Bryson, T.D., Jacobsen, S.E. and Henikoff, S. (2005) Chromatin and siRNA pathways cooperate to maintain DNA methylation of small transposable elements in *Arabidopsis*. *Genome Biol.* 6, R90.
- Van den Broeck, D., Maes, T., Sauer, M., Zethof, J., De Keuleleire, P., D'Hauw, M., Van Montagu, M. and Gerats, T. (1998) *Transposon display* identifies individual transposable elements in high copy number lines. *Plant J.* 13, 121–129.
- Vicient, C.M., Suoniemi, A., Ananthawat-Jónsson, K., Tanskanen, J., Beharav, A., Nevo, E. and Schulman, A.H. (1999) Retrotransposon *BARE-1* and its role in genome evolution in the genus *Hordeum*. *Plant Cell* 11, 1769–1784.
- Vitte, C. and Panaud, O. (2003) Formation of solo-LTRs through unequal homologous recombination counterbalances amplifications of LTR retrotransposons in rice *Oryza sativa* L. *Mol. Biol. Evol.* 20, 528–540.
- Waugh, R., McLean, K., Flavell, A.J., Pearce, S.R., Kumar, A., Thomas, W.T.B. and Powell, W. (1997) Genetic distribution of *BARE-1*-like retrotransposable elements in the barley genome revealed by sequence-specific amplification polymorphisms (S-SAP). *Mol. Gen. Genet.* 253, 687–694.
- Wei, F., Wing, R.A. and Wise, R.P. (2002) Genome dynamics and evolution of the *Mla* (powdery mildew) resistance locus in barley. *Plant Cell* 14, 1903–1917.
- Wicker, T., Stein, N., Albar, L., Feuillet, C., Schlagenhauf, E. and Keller, B. (2001) Analysis of a contiguous 211 kb sequence in diploid wheat (*Triticum monococcum*) reveals multiple mechanisms of genome evolution. *Plant J.* 26, 307–316.
- Wicker, T., Guyot, R., Yahiaoui, N. and Keller, B. (2003) CACTA transposons in Triticeae. A diverse family of high-copy repetitive elements. *Plant Physiol.* 132, 52–63.
- Wicker, T., Sabot, F., Hua-Van, A., Bennetzen, J., Capy, P., Chalhoub, B., Flavell, A.J., Leroy, P., Morgante, M., Panaud, O., Paux, E., SanMiguel, P. and Schulman, A.H. (2007) A unified classification system for eukaryotic transposable elements. *Nat. Rev. Genet.* 8, 973–982.
- Witte, C.P., Le, Q.H., Bureau, T. and Kumar, A. (2001) Terminal-repeat retrotransposons in miniature (TRIM) are involved in restructuring plant genomes. *Proc. Natl. Acad. Sci. USA* 98, 13778–13783.
- Woodhouse, M.R., Freeling, M. and Lisch, D. (2006) Initiation, establishment, and maintenance of heritable MuDR transposon silencing in maize are mediated by distinct factors. *PLoS Biol.* 4, e339.
- Zhang, X., Yazaki, J., Sundaresan, A., Cokus, S., Chan, S.W., Chen, H., Henderson, I.R., Shinn, P., Pellegrini, M., Jacobsen, S.E. and Ecker, J.R. (2006) Genome-wide high-resolution mapping and functional analysis of DNA methylation in *Arabidopsis*. *Cell* 122, 1189–1201.

Chapter 15

Gene and Repetitive Sequence Annotation in the Triticeae

Thomas Wicker and C. Robin Buell

Abstract The Triticeae tribe contains some of the world's most important agricultural crops (wheat, barley and rye) and is perhaps, one of the most challenging for genome annotation because Triticeae genomes are primarily composed of repetitive sequences. Further complicating the challenge is the polyploidy found in wheat and particularly in the hexaploid bread wheat genome. Genomic sequence data are available for the Triticeae in the form of large collections of Expressed Sequence Tags (>1.5 million) and an increasing number of bacterial artificial chromosome clone sequences. Given that high repetitive sequence content in the Triticeae confounds annotation of protein-coding genes, repetitive sequences have been identified, annotated, and collated into public databases. Protein coding genes in the Triticeae are structurally annotated using a combination of *ab initio* gene finders and experimental evidence. Functional annotation of protein coding genes involves assessment of sequence similarity to known proteins, expression evidence, and the presence of domain and motifs. Annotation methods and tools for Triticeae genomic sequences have been adapted from existing plant genome annotation projects and were designed to allow for flexibility of single sequence annotation while allowing a whole community annotation effort to be developed. With the availability of an increasing number of annotated grass genomes, comparative genomics can be exploited to accelerate and enhance the quality of Triticeae sequences annotation. This chapter provides a brief overview of the Triticeae genomes features that are challenging for genome annotation and describes the resources and methods available for sequence annotation with a particular emphasis on problems caused by the repetitive fraction of these genomes.

T. Wicker (✉)

Institute of Plant Biology, University Zurich, Zollikerstrasse 107, CH-8008 Zurich, Switzerland

e-mail: wicker@botinst.uzh.ch

15.1 Triticeae Genomics

Although the Triticeae contains some of the world's most important agricultural crops, this group of plants has only begun to enter the genomics era. This is not due to a lack of interest or need for genomics of Triticeae species but results from the technical challenges of obtaining the genomic sequence from large, repetitive and sometimes polyploid species. The genome of hexaploid, or bread wheat (*Triticum aestivum* L., $2N=6X=42$), is reported to be 16 Gb (Arumuganathan and Earle 1991) and to contain more than 90% repetitive sequences (Li et al. 2004) thereby presenting limitations, primarily fiscal in nature, to current sequencing methodologies. Barley (*Hordeum vulgare* L., $2N=2X=14$) is diploid and has a genome size comparable to that of diploid wheat (5.7 Gb (Bennett and Smith 1976)) with a similar content of repetitive DNA (Smith and Flavell 1975). In the past decade, however, the development of new genomic resources such as bacterial artificial chromosome (BAC) libraries and large collections of markers including Expressed Sequence Tags (ESTs; see Chapter 9) have allowed the establishment of robust genomics programs in the Triticeae including a wheat and a barley genome sequencing initiative (see Chapter 24). The International Wheat Genome Sequencing Consortium (IWGSC; <http://www.wheatgenome.org/>) is focusing its effort on hexaploid wheat, specifically the cultivar Chinese Spring (Gill et al. 2004), and members of the initiative have already generated a number of resources (BAC libraries, physical maps) that allow targeted genome sequencing. The International Barley Sequencing Consortium (IBSC; <http://barleygenome.org>) was also launched to develop genomic resources for genome sequencing of cultivar Morex. For more details on the genome sequencing initiatives for the Triticeae see Chapter 24.

The ideal outcome of a whole genome annotation effort would be a set of genes accurately identified with information about their location on linkage maps and their putative functions. Ancillary annotations such as expression patterns, promoter sequences, orthologous and paralogous sequences are also informative for biologists, breeders, and geneticists but they are not part of a "core" genome annotation. The foundation of an annotation project is the accurate identification of protein coding genes. This is obtained through a combination of computational predictions such as *ab initio* gene finders and through experimental evidence such as transcripts and protein alignments. Appropriate weighting these data types and constructing accurate gene models for an entire genome is generally extremely challenging and becomes a major issue for genomes as complex and large as those of the Triticeae. Thus, successful genome annotation projects result in different gene subsets ranging from well annotated genes (i.e., genes with full length cDNA support) to reasonably annotated (i.e., genes with EST and/or protein support) and genes annotated with less confidence (i.e., genes predicted solely by an *ab initio* gene finder). With respect to functional annotation of large genomes, putative function is

primarily assigned through sequence similarity with other sequenced genomes which is highly prone to transitive annotation errors. This can be addressed by manual curation or through annotation of functional domains such as Pfam domains (Finn et al. 2006) rather relying on “best hits” to a large non-redundant amino acid database of primarily uncured entries (UniProt Consortium 2007).

The high repetitive sequence content of the Triticeae genomes complicates the annotation process in two ways: First because of their abundance it contributes in a significant manner to the bulk of sequence that needs to be processed during the annotation phase and second because some of the repetitive elements are expressed and have features of protein coding genes that can confound gene annotation efforts. Efforts have already begun to address these challenges by developing adequate and efficient bioinformatics tools and resources for interpretation of the Triticeae genome sequences and to ensure accessibility to the scientific community.

15.2 Triticeae Genome Sequence and Annotation Data

15.2.1 *The Triticeae Transcriptome*

ESTs provide a rapid form of gene discovery as they represent the genic portions of the genomes thereby bypassing the large tracts of genome sequence that does not encode for RNA or proteins. They provide a mechanism for gene discovery for species with large and unsequenced genomes such as those of the Triticeae. In 1998, the Triticeae community established an international collaborative network, the International Triticeae EST Cooperative (ITEC, <http://wheat.pw.usda.gov/genome/>) to produce large collections of ESTs. To date, >1.5 million ESTs for wheat, barley, and rye are present in the EST database at NCBI (<http://www.ncbi.nlm.nih.gov/dbEST/index.html>). The August 8, 2008 release of dbEST (080808) contained 1,050,791 ESTs from *Triticum aestivum* (bread wheat), 478,734 ESTs from *Hordeum vulgare* and subsp. *vulgare* (barley), 24,161 from *Hordeum vulgare* subsp. *spontaneum*, 17,381 ESTs from *Triticum turgidum* subsp. *durum* (durum wheat), 11,190 ESTs from *Triticum monococcum*, 9,298 from *Secale cereale* (rye), 1,938 ESTs from *Triticum turgidum*, and 4,315 ESTs from *Aegilops speltoides*. As these ESTs are primarily derived from non-normalized cDNA libraries, redundancy is rampant making them difficult to work with on an individual basis. Thus, these ESTs, along with cloned mRNAs and cDNAs, are typically clustered and assembled into a smaller, representative set of transcripts (unigenes, transcript assemblies, tentative consensus sequences) prior to their use by biologists or bioinformaticians. A number of laboratories provide these clustered assemblies as part of their resource efforts including PlantGDB (<http://www.plantgdb.org/prj/ESTCluster/index.php>), Dana Farber Gene Indices (<http://compbio.dfci>

harvard.edu/tgi/), the TIGR Plant Transcript Assemblies (<http://plantta.tigr.org/>), HarvEST (<http://harvest.ucr.edu/>) and GenoplanteDB (<http://urgi.versailles.inra.fr/GnpSeq/>). Gramene (<http://gramene.org>), a major database for research on grasses, has also mapped EST assemblies to the rice and maize genomes.

The availability of the Triticeae EST collections has allowed for studies on the Triticeae transcriptome (Chao et al. 2006; Houde et al. 2006; Kawaura et al. 2005; Laudencia-Chingcuanco et al. 2006; Mochida et al. 2006; Ogihara et al. 2004), the development of bin-mapped markers for wheat genetic mapping (Conley et al. 2004; Hossain et al. 2004; Lazo et al. 2004; Linkiewicz et al. 2004; Miftahudin et al. 2004; Munkvold et al. 2004; Peng et al. 2004; Qi et al. 2004; Randhawa et al. 2004), EST maps in barley (Stein et al. 2007), and comparative studies of the syntenic relationships between wheat and rice (Conley et al. 2004; Francki et al. 2004; La Rota and Sorrells 2004; Linkiewicz et al. 2004; Peng et al. 2004; Salse et al. 2008; See et al. 2006; Sorrells et al. 2003).

ESTs, along with full-length cDNA clones, are valuable not only for gene discovery but also for empirical evidence that can be used in structural annotation of genomic sequences. The optimal transcript resource is a set of full length (FL) cDNA sequences that provide a complete representation of the full 5' and 3' untranslated regions and the precise location of intron/exon splice junctions for an unambiguous annotation of the gene structure. They have been instrumental in the annotation of genome sequences, including Arabidopsis and rice (Castelli et al. 2004; Haas et al. 2003; Ohyanagi et al. 2006; Ouyang et al. 2007; Tanaka et al. 2008). In addition to their use in structural annotation, the cDNA clone from which the EST or FL cDNA sequence is derived is highly desirable as a resource for functional genomics studies such as overexpression studies. Several thousands of FL cDNA sequences are available for wheat and barley. A query of Genbank (August 17, 2008) revealed 2,033 and 5,510 FL cDNA sequences for wheat and barley, respectively. A project is also in progress to generate FL cDNAs for Chinese Spring, the hexaploid wheat cultivar selected for genome sequencing by the IWGSC (see Chapter 24). To date, ~4,200 FL cDNA sequences have been produced (http://wheat.pw.usda.gov/ITMI/ITMI2005_Proceedings/Abstracts/Ogihara.html). A similar project is in progress for barley (<http://www.shigen.nig.ac.jp/barley/>).

15.2.2 The Triticeae Genomes

Both the wheat and barley communities are pursuing BAC-based sequencing initiatives to obtain the genome sequence (see Chapter 24) after physical maps have been established (see Chapter 11). In the near future, the sequence of ~200 BACs randomly selected from Chinese Spring, will be made available as part of a survey of the wheat genome landscape (Devos et al. 2005). Targeted sequencing of wheat chromosome or chromosome arm specific BAC libraries

are underway including chromosome 3B (Gill et al. 2004; Paux et al. 2006) (http://www.international.inra.fr/research/some_examples/sequencing_the_wheat_genome) and 3AS (<http://wheat.tigr.org/tdb/e2k1/tae1/>). A continually updated list of bread wheat sequencing activities can be seen on the IWGSC web page (<http://www.wheatgenome.org/>). These random as well as targeted sequencing projects will provide ample sequence for optimization and improvement of genome annotation tools. Indeed, to date, genome sequence (excluding ESTs) is available for 67.8 Mb of *Triticum* species in Genbank including 30.3 Mb from the HTG division (draft sequences of BACs), 27.1 Mb from the GSS division (single pass sequences, typically end sequences of BACs), and 10.4 Mb from the PLN division (finished sequence). For *Hordeum* species, 21.8 Mb of genome sequence (excluding ESTs) is available in Genbank including 216.5 Kb from the HTG division, 1.3 Mb from the GSS division, and 20.3 Mb from the PLN division.

15.2.3 Genome Annotation: Structural and Functional Annotation

In the framework of the IWGSC, a working group of Triticeae biologists and bioinformaticians has been established to set up guidelines and develop a community effort for annotating Triticeae genomic sequences. The guidelines are focused on establishing a minimum set of annotations and processes that are provided to the research community for accurate, homogeneous and insightful interpretation of the sequence. The current foci of the guidelines are structural annotation and to a smaller extent, functional annotation. The guidelines, summarized below, are available at the IWGSC web site (<http://www.wheatgenome.org/tool.html>).

The starting point for annotating Triticeae genomic sequences is the identification and annotation of repetitive sequences that compose most of the Triticeae genomes (>80%). Their composition and identification are described in the section below. Following identification, the repetitive sequences are then “masked” to prevent them from confounding annotation of protein coding genes. Genes are identified in the repeat-masked sequence using *ab initio* gene finders. Although multiple gene finders can be used, at a minimum, FGENESH (monocot matrix; Salamov and Solovyev 2000) must be run on the sequence. The sensitivity and specificity of various gene finders on wheat sequences have not been compared and documented although anecdotal evidence suggests that FGENESH is the most accurate *ab initio* gene finder currently available. Gene structure can be improved using transcript and protein evidence to construct an improved gene model. Nomenclature of the transcriptional unit and loci are outlined in the IWGSC annotation guidelines. Standardization of the nomenclature, even at the early stages of a genome effort, is essential to minimizing population of databases with genes, gene models, and transcripts with divergent annotations.

Putative function for the protein encoded in a gene model is determined based on either the presence of a Pfam domain or through sequence similarity evidence. A gene model can be annotated as encoding a “known”, “putative”, “XX-domain containing”, “expressed”, “conserved hypothetical” or “hypothetical” protein depending on the extent of sequence similarity detected. For annotating a gene model as encoding a “known” protein, high sequence similarity (>90–100% identity and coverage) to a characterized protein within an amino acid database such as UniProt (Suzek et al. 2007) must be detected. Expression evidence in the form of alignment to an EST, cDNA or mRNA is optional annotation for the gene model. When a lower level of similarity with an entry in an amino acid database (>45% identity, >50% coverage), and thereby a lower confidence, is observed, the gene model is annotated as encoding a “putative XX” protein. Again, expression evidence is an optional yet informative layer of annotation. For gene models encoding proteins that lack similarity to an entry in an amino acid database but have a Pfam domain above the trusted cutoff, the gene model is annotated as encoding a “XX-domain containing protein”. Here, although expression evidence is optional, its availability is highly informative for deducing the function of the gene model. Sometimes, gene models can have strong sequence similarity with proteins in the amino acid database without known function. In this case, they are referred to as expressed or hypothetical genes. Triticeae genes that match such an entry (>45% identity, >50% coverage) and lack sequence similarity with Triticeae ESTs, mRNAs, or cDNAs (<95% ID, <70% length), are annotated as encoding “conserved hypothetical proteins”. For the gene models that lack substantial sequence similarity (>45% identity, >50% coverage) with a known or putative protein entry in an amino acid database as well as a Pfam domain above the trusted cutoff, but have sequence identity to an EST, cDNA, or mRNA (>95% ID, >70% length), the gene model is annotated as encoding an “expressed protein”. When a gene model lacks any sequence similarity (>45% identity, >50% coverage) with an entry in an amino acid database or with an EST, cDNA or mRNA, the gene model is annotated as encoding a “hypothetical protein”. The availability of transcript support is highly valuable as this is empirical evidence that the gene is transcribed thereby providing more confident annotation than that of hypothetical gene models.

According to the guidelines, additional annotations should be made for Triticeae genes. For example, the top match to the predicted rice and Arabidopsis proteomes should be provided. The rationale behind this is to provide links to well characterized plant genomes in which not only a complete genome sequence and genome annotation datasets are available, but functional resources and data are available to test hypotheses regarding the function of the Triticeae homolog.

Annotation can be done manually, semi-automatically, or automatically. A large factor in determining the approach is the available manpower and the level of quality of annotation desired. Certainly, manual annotation provides a

high quality of interpretation as individual evidence can be weighted and new data from the literature or expert knowledge can be evaluated and incorporated on an ad hoc basis. However, manual annotation is very time consuming and cannot be envisaged for the Triticeae genomes. Thus, the majority of annotation for the wheat and barley genomes will be automated or semi-automated. This is similar to trends in a number of plant genome projects in which the genome has been annotated using automated and semi-automated methods with targeted curation of genes and gene families (Jaillon et al. 2007; Ouyang et al. 2007; Tuskan et al. 2006). For wheat, a semi-automated publicly available annotation pipeline, the TriAnnot pipeline (<http://urgi.versailles.inra.fr/projects/TriAnnot/>) has been established and proposed for the semi-automated annotation of Triticeae genomic sequences. It proposes annotation of wheat and barley BAC sequences through gene and transposon prediction and modeling. Through a simple online process, users can submit their BAC for annotation by the TriAnnot pipeline. Annotation output is provided in a number of formats for downstream analysis including editing in graphical viewers.

Clearly, annotation of Triticeae genomic sequences is in its infancy. As more genome sequence becomes available, training sets (genomic DNA and cognate full length cDNA sequences) will be available allowing for training and improvement of *ab initio* gene finders. Consequently, better characterization and cataloguing of Triticeae repetitive elements will allow for refinement of the gene space and reduce contamination of the gene complement with transposable elements (TEs). However, perhaps the greatest improvement in wheat genome annotation will be from comparative alignments with genome sequences from other Poaceae species. In addition, all annotation is iterative in nature and even for Arabidopsis, in which all of the genes were manually curated (Arabidopsis Genome Initiative 2000), the genome is continually re-annotated as new evidence types and computational methods become available ((Haas et al. 2005), <http://arabidopsis.org/>). Thus, with the large size of the wheat and barley genomes, it will be important that efficient automated/semi-automated annotation pipelines are established which can handle the large sequences such as pseudomolecules as optimal annotation is performed on a genome-scale, not on a small representation scale such as that of a BAC (~120 kb).

15.2.4 Comparative Genome Annotation

For genes, comparative genome annotation is a powerful tool because it can highlight conserved and diverged features among genomes. The availability of the complete rice genome sequence along with genomic sequences of several hundred kb from wheat and barley has allowed comparison between these genomes at the sequence level providing data on the degree of conservation between the grass genomes (Bossolini et al. 2007; Dubcovsky et al. 2001;

Table 15.1 List of Poaceae species with genome sequence and/or pending genome sequence

Species	Tribe	References
<i>Oryza sativa</i> (rice)	Ehrhartoideae	IRGSP (The International Rice Genome Sequencing Project 2005).
<i>Zea mays</i> (maize)	Panicoideae	http://www.maizesequence.org
<i>Brachypodium distachyon</i>	Pooideae	http://www.jgi.doe.gov/sequencing/why/CSP2007/brachypodium.html
<i>Sorghum bicolor</i>	Panicoideae	http://www.phytozome.net/sorghum
<i>Setaria italica</i> (foxtail millet)	Panicoideae	http://jgi.doe.gov/sequencing/why/CSP2008/foxtailmillet.html

Griffiths et al. 2006) (see Chapter 17). In the next few years, the sequence of multiple Poaceae species will be available (Table 15.1) and this will provide important resources for improving genome annotation in this family. This has already been seen in annotation of the rice genome (Zhu and Buell 2007). The use of comparative alignments between rice, maize and sorghum provide information that can (1) improve the structural annotation due to sequence conservation of coding regions, (2) increase confidence of gene predictions in which no transcript support is available, and (3) provide new evidence for functional annotation inferences which can be drawn from experimental and literature reports between orthologous genes.

In a comparison of *Brachypodium* with rice (Bossolini et al. 2007), the annotation of both rice and *Brachypodium* could be improved through such comparative analyses. For rice, seven of the 47 annotated genes could be updated in their structure based on comparative alignments with the collinear *Brachypodium* sequence (Bossolini et al. 2007). Recent reports of comparative analyses between wheat and *Brachypodium* have confirmed the close relationship between genes, gene structure and gene order within the Pooideae (Bossolini et al. 2007; Griffiths et al. 2006). The availability of the *Brachypodium* sequence in the near future (see Chapter 16), will greatly facilitate efforts in understanding the Triticeae genomes structure and composition.

15.3 Repetitive Sequences in the Triticeae

15.3.1 Methods for the Identification of Transposable Elements

The easiest way to identify TEs is by a BLAST (Basic Local Alignment Search Tool) search of the sequence of interest against a database containing known TE sequences. BLAST searches can be done at the DNA (BLASTN) or protein level (BLASTX). BLASTN helps identify closely related TEs which belong to the same family. Usually, the entire element (coding and non-coding parts) can be detected that way. If a TE is more divergent and does not belong to a family already present in the databases, a BLASTX search can help identify protein

coding regions and thus allow determination to which superfamily the TE belongs. The non-coding portions of the TE cannot be characterized by BLASTX and other methods have to be used to determine the exact borders of the element (see below). The ability to identify TEs by BLAST entirely depends on the completeness of the TE database. Whenever a novel repeat is present on the sequence, it will remain undetected.

De novo detection of repeats is more labour intensive and requires a great expertise in the structure and characteristics of repeats. However, it is an important process because as soon as one member of a family is newly identified and characterized, it can be added to the existing databases and further copies of that family can then be identified by sequence comparisons. *De novo* detection is mainly done by searching for coding sequences that are similar to those of known TEs and identification of terminal repeat sequences. Coding sequences are again identified by BLASTX against a series of databases which can (and should) also include animal, fungal, and bacterial sequences.

As mentioned in Chapter 14 on the genomics of TEs, many transposable elements are non-autonomous and do not contain any coding regions. It is also

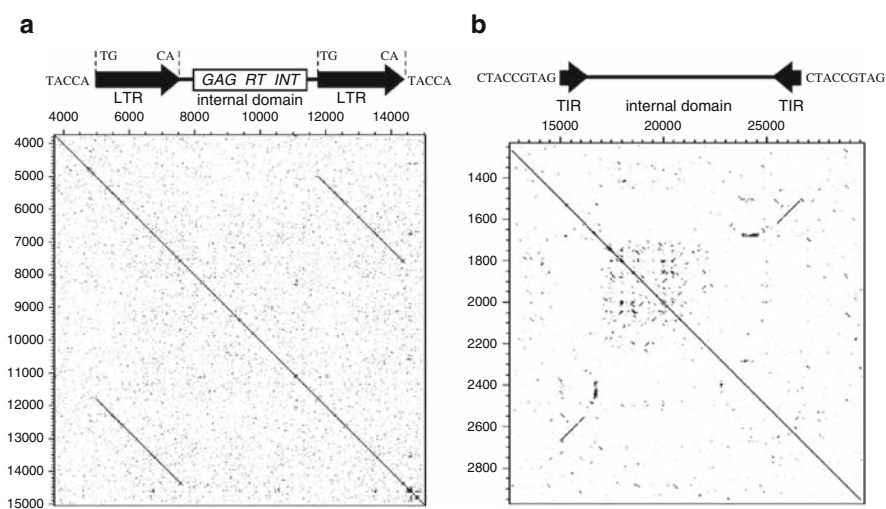


Fig. 15.1 Identification of TEs based on structural characteristics. The structure of the TE is displayed above a DotPlot in which the sequence containing the TE is aligned against itself to visualize repeat structures such as long terminal repeats (LTRs) or terminal inverted repeats (TIRs). A DotPlot is a visual alignment of two sequences, one horizontally and one vertically (the case illustrated here correspond to the alignment of a sequences against itself). The full *diagonal line* from the *top left* to the *bottom right* is the 100% match of the sequence on itself. Other *diagonal lines* represent repeat structures. Direct repeats (LTRs) are parallel to the main *diagonal line* while inverted repeats (TIRs) are perpendicular to it. Other diagnostic features such as canonical LTR termini and target site duplications (TSD) are also easy to detect on such representation (with zooming possibilities on specific regions). **(a)** DotPlot and characteristics of a *BARE1* LTR retrotransposon (5 bp TSD). **(b)** DotPlot of a Mutator transposon with no coding capacity (9 bp TSD)

possible that a new TE family contains highly divergent coding regions that can not be identified based on homology to known elements. In such cases, the TE has to be identified based on structural characteristics such as their terminal repeat sequences (reviewed by (Wicker et al. 2007)). An efficient tool for this task is a so-called DotPlot (Fig. 15.1) which aligns two sequences graphically, one on the x -axis and one on the y -axis. Whenever there is a short stretch of homology (e.g. 5 bp), the programs produces a dot at this position, allowing to easily identify long regions of homology. If a sequence is aligned with itself, DotPlot can be used to identify repeat structures within that sequence (e.g. terminal repeats of TEs).

Almost all TE superfamilies create a so-called target site duplication (TSD) when they insert into the genome (Fig. 15.2). The TSD (also called a “genetic footprint”) is created because the Integrase or Transposase enzymes usually produce staggered ends with overhangs of 2–10 bp.

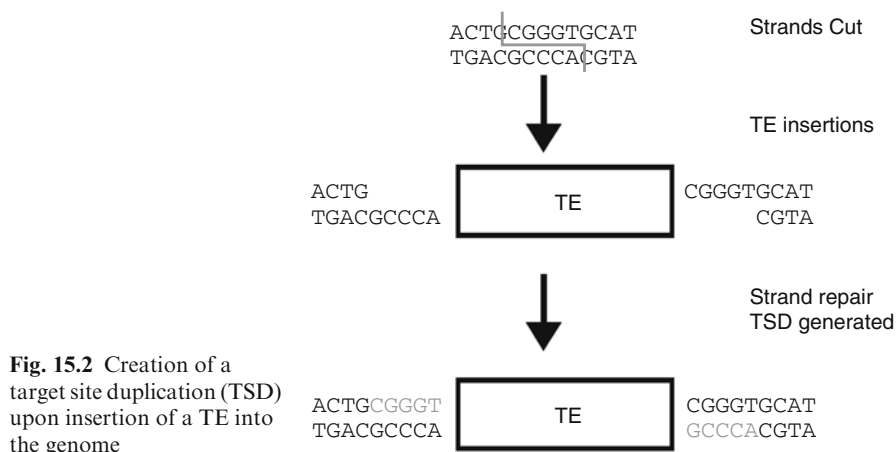


Fig. 15.2 Creation of a target site duplication (TSD) upon insertion of a TE into the genome

In the case of LTR retrotransposons, one would search for the presence of direct repeats that are separated from each other by about 3–5 kb (the usual size of an internal domain) and are flanked by a 5 bp TSD (Fig. 15.1a). Furthermore, LTRs almost always start with TG and end with a CA motif (Fig. 15.1a). DNA transposons can be identified based on their terminal inverted repeats (Fig. 15.1b) as well as the characteristic length of TSDs which usually range from 2–10 bp, depending on the Superfamily (Fig. 15.2).

15.3.2 Problems with Transposable Elements in Triticeae Sequencing

The hundreds of thousands of TE sequences present in the Triticeae genomes have, so far, represented a major barrier to large scale sequencing. To date,

most of the sequencing has been performed in the framework of map-based cloning projects in which a region of usually 2–4 BAC clones is established at the target locus and sequenced by the shotgun-sequencing method. The Sanger sequencing technology which was mostly used until recently only generates sequences of less than 1,000 bp and therefore genomic regions have to be divided into smaller fragments for sequencing. Thus, during shotgun sequencing, the BAC DNA is sheared into small fragments of 3–10 kb which are then sequenced individually. Enough fragments are sequenced to reach a total of 8–10 times the size of the BAC (referred to as 8–10-fold sequencing coverage). The sequenced fragments are then collected together to find overlapping regions, in order to be able to reproduce the original BAC sequence. This process of reconstructing the original sequence is called “sequence assembly”. The product of assembled overlapping sequences is called a “sequence contig”.

The production of the primary (shotgun) sequence itself is not more labour-intensive in the Triticeae than for any other species. Difficulties arise during assembly of the shotgun reads when repetitive sequences are wrongly pooled into artificial contigs and when the sequence of the remaining gaps has to be determined. It is in this phase (called the “finishing” phase) that the TE sequences cause the problems that make Triticeae sequencing so costly and labour-intensive. As of June 2008, there were 377 Triticeae genomic sequences larger than 20 kb available in the NCBI public database (www.ncbi.nlm.nih.gov). Many of them corresponded to individual BAC sequences that were in an unfinished state mostly because of the difficulty to assemble TE regions. If a BAC contains several copies of the same TE, they can cause confusion in the assembly as different copies are assembled into the same sequence contig thereby preventing the correct assembly of the whole sequence. To resolve such mis-assemblies, information from forward and reverse reads of the same shotgun clone can be used and detailed TE annotation of the unfinished sequence can provide hints as to the correct linear order of the sequence contigs. Often, the two LTRs of a LTR-retrotransposon cause the same effect as they are pooled into one single LTR consensus sequence while the internal domain is assembled into a separate sequence contig with no apparent connection to the rest of the BAC sequence.

Even if they are present in a single copy on the BAC, i.e. they should behave like a normal low-copy sequence, TE can also cause gaps in the BAC sequence because of their sequence composition. For example, the highly repeated TEs of the *BARE1* group (*Angela*, *BARE1* and *WIS*) contain a G/C-rich region within their LTRs that almost in all cases causes sequencing problems. Interestingly, analysis of 16 *Angela* and *WIS* LTRs from several independent BACs showed that the gaps are all found in similar positions and that the region can be narrowed down to a few dozen base pairs that apparently contain the problematic motif (Fig. 15.3a). Similarly, many CACTA transposons contain regions that are very difficult to sequence. Most *Caspar* elements, for example, contain

used for an efficient and quick identification of LTR retrotransposons without requiring a lot of specialized knowledge. The disadvantage of that program is that it does not really take into account the possibility of nested insertions, i.e., TEs inserted into other TEs that are very frequent in the Triticeae genomes. Another example for automated annotation is the program TEnest (Kronmiller and Wise 2008) which identifies TEs based on a search against a TE database and also models their nesting patterns, i.e., the order in which the TEs have inserted into one another. This allows a quick assessment of the genome evolution in a particular locus. However, the main disadvantage of these two programs is that they work only on largely finished sequences. Further automated TE recognition pipelines have been developed (Bao and Eddy 2002; Quesneville et al. 2005). The former is a *de novo* repeat identification software which defines the boundaries of repetitive sequences by multiple sequences alignments of regions that contain particular repeat. The latter employs a “combine evidence” strategy analogous that that used for gene prediction where results from homology based and *de novo* TE identification methods are integrated.

Although such programs are very valuable and helpful tools for sequence analysis, one has to consider their outputs with caution. The automated annotation of TE is very complex and many exceptions and special cases are not handled by the programs because the programmer simply did not know about them at the time of development. A typical example is a deletion that eliminates part of a TE. The computer program might then find the first half of the TE and merge it with the second half of a similar TE further downstream. Such an artifact can cause inconsistencies when the evolution of a locus is being analyzed. Even worse, if a gene is located in between the two merged TEs as it can be interpreted as part of the TE if the results are not checked carefully. Thus, every automated annotation of automatically extracted TE dataset should be inspected carefully if one wants to ensure accurate information about TEs.

15.3.4 The Challenge of the Large Number: Quality in Quantity is Needed

To ensure accurate repeat identification and characterization, it is essential that a high-quality repeat database is available. There are several criteria that define the quality of such a database. A few will be mentioned here:

- The size of the TEs and the structure of their terminal sequence needs to be well identified. This allows exact annotation of the borders of TEs on a given sequence and, thus, efficient making of a considerable fraction of the sequence for further gene identification.
- TEs in the database should not contain nested insertions of other TEs. This can lead to distorted estimates of copy numbers of TEs. If, for example, a low-copy transposon contains an insertion of a miniature inverted repeat transposable element which is present in 10,000 copy

numbers in the genome, a BLAST searches against the TE database will often hit the high-copy element inside the low-copy one. If the BLAST output is not carefully interpreted, one can gain the false impression of the abundance of the low-copy element.

- TEs in the databases should not contain genes or fragments thereof. Especially when the TE dataset is produced automatically, as described above, there is the danger that it contains artifact TEs which contain genes or gene fragments.
- TEs are often wrongly annotated as genes, since they may contain coding sequences which are not clearly homologous to typical TE proteins such as Transposase or Reverse Transcriptase. Once a TE is wrongly labeled as a gene, the mistake will continue to be carried on, as future researchers, who come across that particular TE will annotate it again as a gene. This can result in potentially large artifactual “gene families”.

A number of TE databases have been created over the years, with RepBase being the pioneer (Jurka 2000). Several TE databases for plants have been generated as a result of the complete sequencing of the rice and Arabidopsis genomes. The only database dedicated to Triticeae is TREP (Triticeae repeat database, <http://wheat.pw.usda.gov/ITMI/Repeats/>). The most recent release contained over 1,400 TEs sequences representing 180 families. Considering the small set of sequences that is publicly available and the vast size of the Triticeae genome, one has to expect that there are thousands of different TE families yet to be discovered. Classification and annotation of such a large number of TEs can only be precise and reliable if a high quality of the repeat database is maintained even when the number of TEs reaches tens of thousands. So far, the TREP database was curated by a very small number of people, thus, providing relatively consistency in quality. However, the challenges that lie ahead will require the definition of clear guidelines and quality control to provide a system for many dozens or even hundreds of researchers. The first steps were taken by creation of the IWGSC annotation guideline and a proposal for a unified classification system for transposable elements (Wicker et al. 2007).

Acknowledgments Research in the Buell lab on wheat genomics is supported by the National Research Initiative (NRI) Plant Genome Program of the USDA Cooperative State Research, Education and Extension Service (CSREES).

References

- Arabidopsis Genome Initiative. (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408, 796–815.
- Arumuganathan, K. and Earle, E. (1991) Nuclear DNA content of some important plant species. *Plant Mol. Biol. Rep.* 9, 208–218.
- Bao, Z. and Eddy, S.R. (2002) Automated *de novo* identification of repeat sequence families in sequenced genomes. *Genome Res.* 12, 1269–1276.

- Bennett, M.D. and Smith, J.B. (1976) Nuclear DNA amounts in angiosperms. *Philos. Trans. R Soc. Lond. B Biol. Sci.* 274, 227–274.
- Bossolini, E., Wicker, T., Knobel, P.A. and Keller, B. (2007) Comparison of orthologous loci from small grass genomes *Brachypodium* and rice: implications for wheat genomics and grass genome annotation. *Plant J.* 49, 704–717.
- Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C., Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V., Temple, G., Caboche, M., Weissenbach, J. and Salanoubat, M. (2004) Whole genome sequence comparisons and “full-length” cDNA sequences: a combined approach to evaluate and improve *Arabidopsis* genome annotation. *Genome Res.* 14, 406–413.
- Chao, S., Lazo, G.R., You, F., Crossman, C.C., Hummel, D.D., Lui, N., Laudencia-Chingcuanco, D., Anderson, J.A., Close, T.J., Dubcovsky, J., Gill, B.S., Gill, K.S., Gustafson, J.P., Kianian, S.F., Lapitan, N.L., Nguyen, H.T., Sorrells, M.E., McGuire, P.E., Qualset, C.O. and Anderson, O.D. (2006) Use of a large-scale Triticeae expressed sequence tag resource to reveal gene expression profiles in hexaploid wheat (*Triticum aestivum* L.). *Genome* 49, 531–544.
- Conley, E.J., Nduati, V., Gonzalez-Hernandez, J.L., Mesfin, A., Trudeau-Spanjers, M., Chao, S., Lazo, G.R., Hummel, D.D., Anderson, O.D., Qi, L.L., Gill, B.S., Echalié, B., Linkiewicz, A.M., Dubcovsky, J., Akhunov, E.D., Dvorak, J., Peng, J.H., Lapitan, N.L., Pathan, M.S., Nguyen, H.T., Ma, X.F., Miftahudin, Gustafson, J.P., Greene, R.A., Sorrells, M.E., Hossain, K.G., Kalavacharla, V., Kianian, S.F., Sidhu, D., Dilbirligi, M., Gill, K.S., Choi, D.W., Fenton, R.D., Close, T.J., McGuire, P.E., Qualset, C.O. and Anderson, J.A. (2004) A 2600-locus chromosome bin map of wheat homoeologous group 2 reveals interstitial gene-rich islands and colinearity with rice. *Genetics* 168, 625–637.
- Devos, K.M., Ma, J., Pontaroli, A.C., Pratt, L.H. and Bennetzen, J.L. (2005) Analysis and mapping of randomly chosen bacterial artificial chromosome clones from hexaploid bread wheat. *Proc. Natl. Acad. Sci. USA* 102, 19243–19248.
- Dubcovsky, J., Ramakrishna, W., SanMiguel, P.J., Busso, C.S., Yan, L., Shiloff, B.A. and Bennetzen, J.L. (2001) Comparative sequence analysis of colinear barley and rice bacterial artificial chromosomes. *Plant Physiol.* 125, 1342–1353.
- Finn, R.D., Mistry, J., Schuster-Bockler, B., Griffiths-Jones, S., Hollich, V., Lassmann, T., Moxon, S., Marshall, M., Khanna, A., Durbin, R., Eddy, S.R., Sonnhammer, E.L. and Bateman, A. (2006) Pfam: clans, web tools and services. *Nucl. Acids Res.* 34, D247–D251.
- Francki, M., Carter, M., Ryan, K., Hunter, A., Bellgard, M. and Appels, R. (2004) Comparative organization of wheat homoeologous group 3S and 7L using wheat-rice synteny and identification of potential markers for genes controlling xanthophyll content in wheat. *Funct. Integr. Genomics* 4, 118–130.
- Gill, B.S., Appels, R., Botha-Oberholster, A.M., Buell, C.R., Bennetzen, J.L., Chalhou, B., Chumley, F., Dvorak, J., Iwanaga, M., Keller, B., Li, W., McCombie, W.R., Ogihara, Y., Quetier, F. and Sasaki, T. (2004) A workshop report on wheat genome sequencing: International Genome Research on Wheat Consortium. *Genetics* 168, 1087–1096.
- Griffiths, S., Sharp, R., Foote, T.N., Bertin, I., Wanous, M., Reader, S., Colas, I. and Moore, G. (2006) Molecular characterization of Ph1 as a major chromosome pairing locus in polyploid wheat. *Nature* 439, 749–752.
- Haas, B.J., Delcher, A.L., Mount, S.M., Wortman, J.R., Smith, R.K., Jr., Hannick, L.I., Maiti, R., Ronning, C.M., Rusch, D.B., Town, C.D., Salzberg, S.L. and White, O. (2003) Improving the *Arabidopsis* genome annotation using maximal transcript alignment assemblies. *Nucl. Acids Res.* 31, 5654–5666.
- Haas, B.J., Wortman, J.R., Ronning, C.M., Hannick, L.I., Smith, R.K., Jr., Maiti, R., Chan, A.P., Yu, C., Farzad, M., Wu, D., White, O. and Town, C.D. (2005) Complete reannotation of the *Arabidopsis* genome: methods, tools, protocols and the final release. *BMC Biol.* 3, 7.

- Hossain, K.G., Kalavacharla, V., Lazo, G.R., Hegstad, J., Wentz, M.J., Kianian, P.M., Simons, K., Gehlhar, S., Rust, J.L., Syamala, R.R., Obeori, K., Bhamidimarri, S., Karunadharm, P., Chao, S., Anderson, O.D., Qi, L.L., Echaliier, B., Gill, B.S., Linkiewicz, A.M., Ratnasiri, A., Dubcovsky, J., Akhunov, E.D., Dvorak, J., Miftahudin, Ross, K., Gustafson, J.P., Radhawa, H.S., Dilbirligi, M., Gill, K.S., Peng, J.H., Lapitan, N.L., Greene, R.A., Bermudez-Kandianis, C.E., Sorrells, M.E., Feril, O., Pathan, M.S., Nguyen, H.T., Gonzalez-Hernandez, J.L., Conley, E.J., Anderson, J.A., Choi, D.W., Fenton, D., Close, T.J., McGuire, P.E., Qualset, C.O. and Kianian, S.F. (2004) A chromosome bin map of 2148 expressed sequence tag loci of wheat homoeologous group 7. *Genetics* 168, 687–699.
- Houde, M., Belcaid, M., Ouellet, F., Danyluk, J., Monroy, A.F., Dryanova, A., Gulick, P., Bergeron, A., Laroche, A., Links, M.G., MacCarthy, L., Crosby, W.L. and Sarhan, F. (2006) Wheat, E.S.T resources for functional genomics of abiotic stress. *BMC Genomics* 7, 149.
- Jaillon, O., Aury, J.M., Noel, B., Policriti, A., Clepet, C., Casagrande, A., Choisne, N., Aubourg, S., Vitulo, N., Jubin, C., Vezzi, A., Legeai, F., Hugueney, P., Dasilva, C., Horner, D., Mica, E., Jublot, D., Poulain, J., Bruyere, C., Billault, A., Segurens, B., Gouyvenoux, M., Ugarte, E., Cattonaro, F., Anthouard, V., Vico, V., Del Fabbro, C., Alaux, M., Di Gaspero, G., Dumas, V., Felice, N., Paillard, S., Juman, I., Moroldo, M., Scalabrin, S., Canaguier, A., Le Clainche, I., Malacrida, G., Durand, E., Pesole, G., Laucou, V., Chatelet, P., Merdinoglu, D., Delledonne, M., Pezzotti, M., Lecharny, A., Scarpelli, C., Artiguenave, F., Pe, M.E., Valle, G., Morgante, M., Caboche, M., Adam-Blondon, A.F., Weissenbach, J., Quetier, F. and Wincker, P. (2007) The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nature* 449, 463–467.
- Jurka, J. (2000) Repbase update: a database and an electronic journal of repetitive elements. *Trends Genet.* 16, 418–420.
- Kawaura, K., Mochida, K. and Ogiwara, Y. (2005) Expression profile of two storage-protein gene families in hexaploid wheat revealed by large-scale analysis of expressed sequence tags. *Plant Physiol.* 139, 1870–1880.
- Kronmiller, B.A. and Wise, R.P. (2008) TEnest: automated chronological annotation and visualization of nested plant transposable elements. *Plant Physiol.* 146, 45–59.
- La Rota, M. and Sorrells, M.E. (2004) Comparative DNA sequence analysis of mapped wheat ESTs reveals the complexity of genome relationships between rice and wheat. *Funct. Integr. Genomics* 4, 34–46.
- Laudencia-Chinguanco, D.L., Stamova, B.S., Lazo, G.R., Cui, X. and Anderson, O.D. (2006) Analysis of the wheat endosperm transcriptome. *J. Appl. Genet.* 47, 287–302.
- Lazo, G.R., Chao, S., Hummel, D.D., Edwards, H., Crossman, C.C., Lui, N., Matthews, D. E., Carollo, V.L., Hane, D.L., You, F.M., Butler, G.E., Miller, R.E., Close, T.J., Peng, J. H., Lapitan, N.L., Gustafson, J.P., Qi, L.L., Echaliier, B., Gill, B.S., Dilbirligi, M., Randhawa, H.S., Gill, K.S., Greene, R.A., Sorrells, M.E., Akhunov, E.D., Dvorak, J., Linkiewicz, A.M., Dubcovsky, J., Hossain, K.G., Kalavacharla, V., Kianian, S.F., Mahmoud, A.A., Miftahudin, Ma, X.F., Conley, E.J., Anderson, J.A., Pathan, M.S., Nguyen, H.T., McGuire, P.E., Qualset, C.O. and Anderson, O.D. (2004) Development of an expressed sequence tag (EST) resource for wheat (*Triticum aestivum* L.): EST generation, unigene analysis, probe selection and bioinformatics for a 16,000-locus bin-delimited map. *Genetics* 168, 585–593.
- Li, W., Zhang, P., Fellers, J.P., Friebe, B. and Gill, B.S. (2004) Sequence composition, organization, and evolution of the core Triticeae genome. *Plant J.* 40, 500–511.
- Linkiewicz, A.M., Qi, L.L., Gill, B.S., Ratnasiri, A., Echaliier, B., Chao, S., Lazo, G.R., Hummel, D.D., Anderson, O.D., Akhunov, E.D., Dvorak, J., Pathan, M.S., Nguyen, H. T., Peng, J.H., Lapitan, N.L., Miftahudin, Gustafson, J.P., La Rota, C.M., Sorrells, M.E., Hossain, K.G., Kalavacharla, V., Kianian, S.F., Sandhu, D., Bondareva, S.N., Gill, K.S.,

- Conley, E.J., Anderson, J.A., Fenton, R.D., Close, T.J., McGuire, P.E., Qualset, C.O. and Dubcovsky, J. (2004) A 2500-locus bin map of wheat homoeologous group 5 provides insights on gene distribution and colinearity with rice. *Genetics* 168, 665–676.
- McCarthy, E.M. and McDonald, J.F. (2003) LTR_STRUC: a novel search and identification program for LTR retrotransposons. *Bioinformatics* 19, 362–367.
- Miftahudin, Ross, K., Ma, X.F., Mahmoud, A.A., Layton, J., Milla, M.A., Chikmawati, T., Ramalingam, J., Feril, O., Pathan, M.S., Momirovic, G.S., Kim, S., Chema, K., Fang, P., Haule, L., Struxness, H., Birkes, J., Yaghoobian, C., Skinner, R., McAllister, J., Nguyen, V., Qi, L.L., Echalié, B., Gill, B.S., Linkiewicz, A.M., Dubcovsky, J., Akhunov, E.D., Dvorak, J., Dilbirligi, M., Gill, K.S., Peng, J.H., Lapitan, N.L., Bermudez-Kandianis, C. E., Sorrells, M.E., Hossain, K.G., Kalavacharla, V., Kianian, S.F., Lazo, G.R., Chao, S., Anderson, O.D., Gonzalez-Hernandez, J., Conley, E.J., Anderson, J.A., Choi, D.W., Fenton, R.D., Close, T.J., McGuire, P.E., Qualset, C.O., Nguyen, H.T. and Gustafson, J.P. (2004) Analysis of expressed sequence tag loci on wheat chromosome group 4. *Genetics* 168, 651–663.
- Mochida, K., Kawaura, K., Shimosaka, E., Kawakami, N., Shin, I.T., Kohara, Y., Yamazaki, Y. and Ogihara, Y. (2006) Tissue expression map of a large number of expressed sequence tags and its application to in silico screening of stress response genes in common wheat. *Mol. Genet. Genomics* 276, 304–312.
- Munkvold, J.D., Greene, R.A., Bermudez-Kandianis, C.E., La Rota, C.M., Edwards, H., Sorrells, S.F., Dake, T., Benscher, D., Kantety, R., Linkiewicz, A.M., Dubcovsky, J., Akhunov, E.D., Dvorak, J., Miftahudin, Gustafson, J.P., Pathan, M.S., Nguyen, H.T., Matthews, D.E., Chao, S., Lazo, G.R., Hummel, D.D., Anderson, O.D., Anderson, J.A., Gonzalez-Hernandez, J.L., Peng, J.H., Lapitan, N., Qi, L.L., Echalié, B., Gill, B.S., Hossain, K.G., Kalavacharla, V., Kianian, S.F., Sandhu, D., Erayman, M., Gill, K.S., McGuire, P.E., Qualset, C.O. and Sorrells, M.E. (2004) Group 3 chromosome bin maps of wheat and their relationship to rice chromosome 1. *Genetics* 168, 639–650.
- Ogihara, Y., Mochida, K., Kawaura, K., Murai, K., Seki, M., Kamiya, A., Shinozaki, K., Carninci, P., Hayashizaki, Y., Shin, I.T., Kohara, Y. and Yamazaki, Y. (2004) Construction of a full-length cDNA library from young spikelets of hexaploid wheat and its characterization by large-scale sequencing of expressed sequence tags. *Genes Genet. Syst.* 79, 227–232.
- Ohyanagi, H., Tanaka, T., Sakai, H., Shigemoto, Y., Yamaguchi, K., Habara, T., Fujii, Y., Antonio, B.A., Nagamura, Y., Imanishi, T., Ikeo, K., Itoh, T., Gojobori, T. and Sakaki, T. (2006) The Rice Annotation Project Database (RAP-DB): hub for *Oryza sativa* ssp. *japonica* genome information. *Nucl. Acids Res.* 34, D741–D744.
- Ouyang, S., Zhu, W., Hamilton, J., Lin, H., Campbell, M., Childs, K., Thibaud-Nissen, F., Malek, R.L., Lee, Y., Zheng, L., Orvis, J., Haas, B., Wortman, J. and Buell, C.R. (2007) The TIGR Rice Genome Annotation Resource: improvements and new features. *Nucl. Acids Res.* 35, D883–D887.
- Paux, E., Roger, D., Badaeva, E., Gay, G., Bernard, M., Sourdille, P. and Feuillet, C. (2006) Characterizing the composition and evolution of homoeologous genomes in hexaploid wheat through BAC-end sequencing on chromosome 3B. *Plant J.* 48, 463–474.
- Peng, J.H., Zadeh, H., Lazo, G.R., Gustafson, J.P., Chao, S., Anderson, O.D., Qi, L.L., Echalié, B., Gill, B.S., Dilbirligi, M., Sandhu, D., Gill, K.S., Greene, R.A., Sorrells, M.E., Akhunov, E.D., Dvorak, J., Linkiewicz, A.M., Dubcovsky, J., Hossain, K.G., Kalavacharla, V., Kianian, S.F., Mahmoud, A.A., Miftahudin, Conley, E.J., Anderson, J.A., Pathan, M.S., Nguyen, H.T., McGuire, P.E., Qualset, C.O., Lapitan, N.L. (2004) Chromosome bin map of expressed sequence tags in homoeologous group 1 of hexaploid wheat and homoeology with rice and Arabidopsis. *Genetics* 168, 609–623.
- Qi, L.L., Echalié, B., Chao, S., Lazo, G.R., Butler, G.E., Anderson, O.D., Akhunov, E.D., Dvorak, J., Linkiewicz, A.M., Ratnasiri, A., Dubcovsky, J., Bermudez-Kandianis, C.E., Greene, R.A., Kantety, R., La Rota, C.M., Munkvold, J.D., Sorrells, S.F., Sorrells, M.E.,

- Dilbirligi, M., Sidhu, D., Erayman, M., Randhawa, H.S., Sandhu, D., Bondareva, S.N., Gill, K.S., Mahmoud, A.A., Ma, X.F., Miftahudin, Gustafson, J.P., Conley, E.J., Nduati, V., Gonzalez-Hernandez, J.L., Anderson, J.A., Peng, J.H., Lapitan, N.L., Hossain, K.G., Kalavacharla, V., Kianian, S.F., Pathan, M.S., Zhang, D.S., Nguyen, H.T., Choi, D.W., Fenton, R.D., Close, T.J., McGuire, P.E., Qualset, C.O. and Gill, B.S. (2004) A chromosome bin map of 16,000 expressed sequence tag loci and distribution of genes among the three genomes of polyploid wheat. *Genetics* 168, 701–712.
- Quesneville, H., Bergman, C.M., Andrieu, O., Autard, D., Nouaud, D., Ashburner, M. and Anxolabehere, D. (2005) Combined evidence annotation of transposable elements in genome sequences. *PLoS Comput. Biol.* 1, 166–175.
- Randhawa, H.S., Dilbirligi, M., Sidhu, D., Erayman, M., Sandhu, D., Bondareva, S., Chao, S., Lazo, G.R., Anderson, O.D., Miftahudin, Gustafson, J.P., Echalié, B., Qi, L.L., Gill, B.S., Akhunov, E.D., Dvorak, J., Linkiewicz, A.M., Ratnasiri, A., Dubcovsky, J., Bermudez-Kandianis, C.E., Greene, R.A., Sorrells, M.E., Conley, E.J., Anderson, J.A., Peng, J.H., Lapitan, N.L., Hossain, K.G., Kalavacharla, V., Kianian, S.F., Pathan, M.S., Nguyen, H.T., Endo, T.R., Close, T.J., McGuire, P.E., Qualset, C.O. and Gill, K.S. (2004) Deletion mapping of homoeologous group 6-specific wheat expressed sequence tags. *Genetics* 168, 677–686.
- Salamov, A.A. and Solovyev, V.V. (2000) *Ab initio* gene finding in *Drosophila* genomic DNA. *Genome Res.* 10, 516–522.
- Salse, J., Bolot, S., Throude, M., Jouffe, V., Piegu, B., Quraishi, U.M., Calcagno, T., Cooke, R., Delseny, M. and Feuillet, C. (2008) Identification and characterization of shared duplications between rice and wheat provide new insight into grass genome evolution. *Plant Cell* 20, 11–24.
- See, D.R., Brooks, S., Nelson, J.C., Brown-Guedira, G., Friebe, B. and Gill, B.S. (2006) Gene evolution at the ends of wheat chromosomes. *Proc. Natl. Acad. Sci. USA* 103, 4162–4167.
- Smith, D.B. and Flavell, R.B. (1975) Characterisation of the wheat genome by renaturation kinetics. *Chromosoma* 50, 223–242.
- Sorrells, M.E., La Rota, M., Bermudez-Kandianis, C.E., Greene, R.A., Kantety, R., Munkvold, J.D., Miftahudin, Mahmoud, A., Ma, X., Gustafson, P.J., Qi, L.L., Echalié, B., Gill, B.S., Matthews, D.E., Lazo, G.R., Chao, S., Anderson, O.D., Edwards, H., Linkiewicz, A.M., Dubcovsky, J., Akhunov, E.D., Dvorak, J., Zhang, D., Nguyen, H.T., Peng, J., Lapitan, N.L., Gonzalez-Hernandez, J.L., Anderson, J.A., Hossain, K., Kalavacharla, V., Kianian, S.F., Choi, D.W., Close, T.J., Dilbirligi, M., Gill, K.S., Steber, C., Walker-Simmons, M.K., McGuire, P.E. and Qualset, C.O. (2003) Comparative DNA sequence analysis of wheat and rice genomes. *Genome Res.* 13, 1818–1827.
- Stein, N., Prasad, M., Scholz, U., Thiel, T., Zhang, H., Wolf, M., Kota, R., Varshney, R.K., Perovic, D., Grosse, I. and Graner, A. (2007) A 1,000-loci transcript map of the barley genome: new anchoring points for integrative grass genomics. *Theor. Appl. Genet.* 114, 823–839.
- Suzek, B.E., Huang, H., McGarvey, P., Mazumder, R. and Wu, C.H. (2007) UniRef: comprehensive and non-redundant UniProt reference clusters. *Bioinformatics* 23, 1282–1288.
- Tanaka, T., Antonio, B.A., Kikuchi, S., Matsumoto, T., Nagamura, Y., Numa, H., Sakai, H., Wu, J., Itoh, T., Sasaki, T., Aono, R., Fujii, Y., Habara, T., Harada, E., Kanno, M., Kawahara, Y., Kawashima, H., Kubooka, H., Matsuya, A., Nakaoka, H., Saichi, N., Sanbonmatsu, R., Sato, Y., Shinso, Y., Suzuki, M., Takeda, J., Tanino, M., Todokoro, F., Yamaguchi, K., Yamamoto, N., Yamasaki, C., Imanishi, T., Okido, T., Tada, M., Ikeo, K., Tateno, Y., Gojobori, T., Lin, Y.C., Wei, F.J., Hsing, Y.I., Zhao, Q., Han, B., Kramer, M.R., McCombie, R.W., Lonsdale, D., O'Donovan, C.C., Whitfield, E.J., Apweiler, R., Koyanagi, K.O., Khurana, J.P., Raghuvanshi, S., Singh, N.K., Tyagi, A.K., Haberer, G., Fujisawa, M., Hosokawa, S., Ito, Y., Ikawa, H., Shibata, M., Yamamoto, M., Bruskiewich, R.M., Hoen, D.R., Bureau, T.E., Namiki, N., Ohyanagi, H., Sakai, Y., Nobushima, S.,

- Sakata, K., Barrero, R.A., Souvorov, A., Smith-White, B., Tatusova, T., An, S., An, G., Oota, S., Fuks, G., Messing, J., Christie, K.R., Lieberherr, D., Kim, H., Zuccolo, A., Wing, R.A., Nobuta, K., Green, P.J., Lu, C., Meyers, B.C., Chaparro, C., Piegu, B., Panaud, O. and Echeverria, M. (2008) The Rice Annotation Project Database (RAP-DB): 2008 update. *Nucl. Acids Res.* 36, D1028–D1033.
- The International Rice Genome Sequencing Project. (2005) The map-based sequence of the rice genome. *Nature* 436, 793–800.
- Tuskan, G.A., Difazio, S., Jansson, S., Bohlmann, J., Grigoriev, I., Hellsten, U., Putnam, N., Ralph, S., Rombauts, S., Salamov, A., Schein, J., Sterck, L., Aerts, A., Bhalerao, R.R., Bhalerao, R.P., Blaudez, D., Boerjan, W., Brun, A., Brunner, A., Busov, V., Campbell, M., Carlson, J., Chalot, M., Chapman, J., Chen, G.L., Cooper, D., Coutinho, P.M., Couturier, J., Covert, S., Cronk, Q., Cunningham, R., Davis, J., Degroove, S., Dejardin, A., Depamphilis, C., Detter, J., Dirks, B., Dubchak, I., Duplessis, S., Ehling, J., Ellis, B., Gendler, K., Goodstein, D., Gribskov, M., Grimwood, J., Groover, A., Gunter, L., Hamberger, B., Heinze, B., Helariutta, Y., Henrissat, B., Holligan, D., Holt, R., Huang, W., Islam-Faridi, N., Jones, S., Jones-Rhoades, M., Jorgensen, R., Joshi, C., Kangasjarvi, J., Karlsson, J., Kelleher, C., Kirkpatrick, R., Kirst, M., Kohler, A., Kalluri, U., Larimer, F., Leebens-Mack, J., Leple, J.C., Locascio, P., Lou, Y., Lucas, S., Martin, F., Montanini, B., Napoli, C., Nelson, D.R., Nelson, C., Nieminen, K., Nilsson, O., Pereda, V., Peter, G., Philippe, R., Pilate, G., Poliakov, A., Razumovskaya, J., Richardson, P., Rinaldi, C., Ritland, K., Rouze, P., Ryaboy, D., Schmutz, J., Schrader, J., Segerman, B., Shin, H., Siddiqui, A., Sterky, F., Terry, A., Tsai, C.J., Uberbacher, E., Unneberg, P., Vahala, J., Wall, K., Wessler, S., Yang, G., Yin, T., Douglas, C., Marra, M., Sandberg, G., Van de Peer, Y. and Rokhsar, D. (2006) The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* 313, 1596–1604.
- UniProt Consortium. (2007) The Universal Protein Resource (UniProt). *Nucl. Acids Res.* 35, D193–D197.
- Wicker, T., Guyot, R., Yahiaoui, N. and Keller, B. (2003) CACTA transposons in Triticeae. A diverse family of high-copy repetitive elements. *Plant Physiol.* 132, 52–63.
- Wicker, T., Sabot, F., Hua-Van, A., Bennetzen, J.L., Capy, P., Chalhoub, B., Flavell, A., Leroy, P., Morgante, M., Panaud, O., Paux, E., SanMiguel, P. and Schulman, A.H. (2007) A unified classification system for eukaryotic transposable elements. *Nat. Rev. Genet.* 8, 973–982.
- Zhu, W. and Buell, C.R. (2007) Improvement of whole-genome annotation of cereals through comparative analyses. *Genome Res.* 17, 299–310.

Chapter 16

Brachypodium distachyon, a New Model for the Triticeae

John Vogel and Jennifer Bragg

Abstract *Brachypodium distachyon* (Brachypodium) is a small annual grass with biological, physical and genomic attributes (e.g. rapid cycling, small stature, inbreeding, small genome, diploid accessions) suitable for use as a modern model system. In pursuit of this goal, researchers have made rapid progress in developing genomic resources that will transform Brachypodium into a powerful model system including: facile *Agrobacterium*-mediated transformation methods, BAC libraries, physical maps, genetic maps, and germ-plasm resources. In addition, a preliminary 4x draft of the entire genome has been released, and completion of the final 8x assembly is anticipated in 2009. This chapter provides an overview of the advantages of Brachypodium as a model system and surveys the use and potential applications of this system to aid wheat, barley and *Lolium* research.

16.1 Model Systems in Biology

Biologists strive for a better understanding of the machinery that drives the living world and the great complexity of biological systems presents a challenge to their endeavors. In response, scientists have sought means to reduce the complexity of their particular systems of study in order to reveal the underlying basic design principles. Although reductionist methods are not appropriate for all experiments, in many areas of biology they are essential to achieve rapid progress. Research using relatively simple model organisms such as *E. coli*, yeast, fruit flies, and mice has led to innumerable discoveries that benefit the daily lives of billions of people. As scientists turn their attention toward more specialized areas of study, the number of model systems continually increases. For example, developmental biology has adopted *C. elegans* and zebrafish as

J. Vogel (✉)

USDA-ARS Western Regional Research Center, Genomics and Gene Discovery Unit,
Albany, CA 94710, USA

e-mail: john.vogel@ars.usda.gov

model organisms. Laboratory manipulations of model organisms are facilitated by simple growth requirements, short generation time, the ability to inbreed, and small size. Additionally, characteristics of model systems that suit them for modern genomic methods include diversity of natural populations, facile transformation, diploidy, simple genetics, and a small genome size.

Plant biologists have widely adopted the small weedy species *Arabidopsis thaliana* as a generalized model plant and, due to its inherent biological attributes and the collegial nature of the research community that has sprung up around it, *Arabidopsis* has become an extremely powerful system. However, *Arabidopsis* is not suitable to study many aspects of grass biology due to the biological differences that have arisen between dicots and monocots in the 150 million years since they last shared a common ancestor. As an example, grass cell walls differ dramatically from dicot cell walls in terms of the major structural polysaccharides present, how those polysaccharides are linked together, and the abundance and importance of pectins, proteins and phenolic compounds (Carpita 1996). A partial list of additional areas in which *Arabidopsis* is not an appropriate model for the study of grasses includes: mycorrhizal associations, architecture of the grass plant, grain properties, intercalary meristems, and grass development. The tremendous importance of grasses as food, feed and, increasingly, as fuel, argue strongly for the development of a truly tractable grass model system. At a first glance, rice with its sequenced genome and large research community would seem to fill this bill. However, upon closer examination, the demanding growing conditions and large size of rice plants make it a poor choice for high-throughput genomic experiments in temperate regions. The fact that rice is a semi-aquatic tropical grass further limits its applicability as a model for temperate grasses, especially in areas like freezing tolerance and vernalization.

16.2 Introduction to *Brachypodium distachyon*

The utility of the small annual grass *Brachypodium distachyon* (hereafter referred to as *Brachypodium*) as a model system for the study of the Triticeae was first discussed in a 2001 paper that pointed out that *Brachypodium* displays all of the biological, physical and genomic attributes required for use as a model system (Draper et al. 2001). The small size and rapid generation time of *Brachypodium* enables high-throughput studies. Densities of 1,000 plants/m² can be easily achieved in growth chambers or greenhouses allowing growth of large numbers of plants under controlled environmental conditions (Fig. 16.1). For comparison, the same space accommodates only 50 wheat plants, 36 rice plants, or four switchgrass plants (Table 16.1). Furthermore, *Brachypodium* is self-fertile and does not typically outcross. This feature is useful for breeding and maintaining homozygous lines for many applications that require the maintenance of large numbers of independent genotypes (e.g. mapping

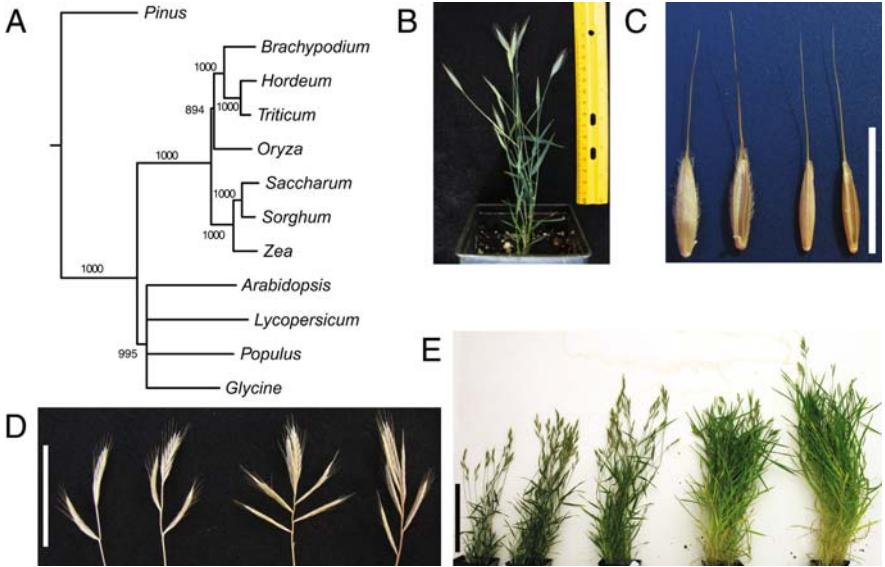


Fig. 16.1 Relationships and phenotypes of *Brachypodium*. (A) Rooted phylogenetic tree based on the combined partial nucleotide sequences of 20 highly expressed genes (Vogel et al. 2006b). Branch length is proportional to sequence divergence. (B) A plant from line Bd21-3 flowering under 20 h light conditions. (C) Close-up of seeds from two accessions. Note that the seeds on the right are hairless and smaller than the seeds on the left. (D) Variation in inflorescence architecture in *four new inbred lines* from Turkey grown under the same conditions. Note the differences in spikelet number and angle. (E) Effect of vernalization on *inbred lines* from one location, Bismil, in Turkey. Plants were placed in the cold for 3 weeks and then moved into a growth chamber with 20 h daylength. The plants in the pot on the left flowered after 25 days in the growth chamber and the plants in the pot on the right flowered after 50 days. Note the range of flowering times even at one location. (D and E) Seeds of Turkish lines were kindly supplied by Metin Tuna (Namik Kemal University, Tekirdag, Turkey). The scale bar in (C) is 1 cm, in (D) is 5 cm and in (E) is 15 cm (See Color Insert)

Table 16.1 Comparison of models and crops

	Brachypodium	Arabidopsis	Rice	Wheat	Switchgrass
Height (cm)	15–20	15–20	100	50	200
Density (plants/m ²)	1,000	2,000	36	50	6
Growth Requirements	simple	simple	demanding	simple	simple
Genome Size (Mbp)	300	165	430	16,000	2,400
Generation Time (weeks)	8–12	8–12	30	12	26
Reproduction	selfing	selfing	selfing	selfing	outcrossing
Cell Wall Type	Type 2	Type 1	Type 2	Type 2	Type 2

experiments, mutant analysis, and studies of natural diversity). As a group, the grasses are notorious for very large genomes. Fortunately, the ~300 Mbp diploid *Brachypodium* genome is one of the smallest of any grass. Since wheat is more closely related to *Brachypodium* than to rice, *Brachypodium* will serve as a more relevant model for wheat structural genomic studies. Like wheat, polyploid *Brachypodium* accessions exist and they may be useful as a model for polyploidy.

16.2.1 Genome Size and Polyploidy

A compact genome is one of the most important attributes of a modern model organism because it permits efficient positional cloning of genes, facilitates genome-wide insertional mutagenesis, and is a prerequisite for whole genome sequencing. Three publications reported the c-value of diploid *Brachypodium* to be approximately 0.36 pg, corresponding to a genome size of approximately 300–320 Mbp (Bennett and Leitch 2005; Vogel and Hill 2008; Vogel et al. 2006a). Although Draper et al. reported a c-value of only 0.15 pg, Bennett and Leitch (2005) obtained a c-value of 0.36 using the same accession, suggesting that the lower number may have been an artifact. An independent estimation of *Brachypodium* genome size was calculated in the range of 300 Mbp based on the frequency of recovering single-copy genes from two bacterial artificial chromosome (BAC) libraries (Huo et al. 2006), and the preliminary 4x genome assembly suggests a size no larger than 300 Mbp (unpublished). Thus, several lines of evidence indicate that *Brachypodium* has one of the smallest genomes of any grass.

Similar to wheat, polyploid *Brachypodium* accessions have been described, and 1n chromosome numbers of 5, 10, and 15 have been reported. Initially, these chromosome numbers appeared to fit a simple autopolyploid series: diploid, tetraploid, and hexaploid. However, careful analysis of c-values and fluorescence *in situ* hybridization (FISH) labeled karyotypes suggests otherwise. FISH analysis of 1n = 10 accessions revealed 10 small chromosomes more similar to the diploid *B. sylvaticum* than to the larger chromosomes found in 1n = 5 accessions (Hasterok et al. 2004). Thus, the 1n = 10 *Brachypodium* accessions appear to be diploid with a 1n = 10 base chromosome number similar to *B. sylvaticum* (Hasterok et al. 2004). The 1n = 15 accessions have c-values approximately twice as large as those of the diploid 1n = 5 accessions (Vogel et al. 2006a) and their karyotypes contain 5 large and 10 small chromosomes (Hasterok et al. 2004). The banding pattern of FISH labeling using a rDNA probe suggests that 1n = 15 accessions are actually allotetraploids with one parent containing a genome similar to diploid 1n = 5 accessions and the other parent containing a genome with a 1n = 10 chromosome number (Hasterok et al. 2004). Taken together, these data suggest that the *Brachypodium* 1n = 10 accessions are actually a separate species and the 1n = 15 accessions are derived

from an interspecific hybrid between one parent similar to the $1n=5$ diploid and one parent similar to the $1n=10$ diploid (Hasterok et al. 2006a,b). The $1n=15$ allotetraploid plants can be easily distinguished from diploid *Brachypodium* ($1n=5$) because the allotetraploids have larger seeds, grow to a larger size, have anthers that typically exert, and do not require vernalization for flowering. In addition, examination of SSR polymorphisms in an allotetraploid $1n=15$ accession revealed multiple bands which presumably correspond to the different genomes (unpublished).

16.2.2 Relationship to Other Grasses

The phylogenetic relationship between the genus *Brachypodium* and the other grasses has been evaluated a number of times with increasing amounts of data. *Brachypodium* has consistently been placed approximately halfway between rice and wheat (Fig. 16.1). Reports based on internal transcribed spacer (ITS) and 5.8 s rDNA sequence (Hsaio et al. 1994), genomic RFLP and RAPD markers (Catalán et al. 1995), and ITS sequence plus the chloroplast *ndfH* gene (Catalán and Olmstead 2000) all placed *Brachypodium* between rice and a clade containing temperate grains like wheat, barley and *Secale*. Additional examinations of a much broader spectrum of grasses used ITS and *ndfH* sequence as well as morphological and chloroplast restriction sites (Kellogg 2001) or the sequence of the *matK* chloroplast gene (Döring et al. 2007). These studies placed *Brachypodium* in the subfamily Pooideae just below the radiation of the small grains and forage and turf grasses making *Brachypodium* “sister” to the temperate grasses of greatest economic significance. However, phylogenies based on single genes or small sets of genes can produce inconsistent phylogenetic trees (Rokas et al. 2003), and this phenomenon as has been observed with rice (Kellogg 1998). Therefore, it was important to examine the phylogenetic relationship of *Brachypodium* using larger datasets. Analysis of a dataset comprising 11 kb of sequence from 20 highly expressed genes verified the relationship between *Brachypodium* and the small grains (Fig. 16.1; Vogel et al. 2006b). An even larger dataset based on 335 BAC end sequences provides further evidence to confirm the placement of *Brachypodium* within the grasses (Huo et al. 2007). Thus, the relationship of *Brachypodium* to cereal crops and other grasses has been firmly established.

16.3 *Brachypodium* as An Experimental System

Brachypodium is easy to grow in large numbers under controlled conditions, is easy to transform, is genetically tractable, and has a small genome. In addition, a large collection of diverse accessions and described inbred lines are currently

being created for *Brachypodium* and will allow researchers to exploit the power of natural diversity for understanding basic questions in biology.

16.3.1 Growth Requirements and Flowering Triggers

One of the strengths of *Brachypodium* as a model system is its ease of culture under laboratory conditions. This is in contrast to rice whose demanding growth requirements, large size, and long generation time are a barrier to many researchers. Generally, *Brachypodium* can be grown in growth chambers or greenhouses used for *Arabidopsis*, wheat or barley. Our standard conditions for growth chambers are: 20 h light: 4 h dark photoperiod, 24°C during the day and 18°C at night with cool-white fluorescent lighting at a level of 150 $\mu\text{Em}^{-2}\text{s}^{-1}$. Our standard greenhouse conditions are: no shading, 24°C in the day and 18°C at night, and supplemental lighting to extend daylength to 16 h. Providing the appropriate conditions to induce flowering is critical to prevent the plants from producing excessive vegetative growth. Vernalization has been shown to induce flowering in all diploid accessions studied to date. However, the time required to induce flowering varies greatly between accessions (Table 16.2). In general, accessions originating from colder northern regions (e.g. northern Turkey) require a longer period of vernalization to induce flowering (e.g. 8–12 weeks), and accessions originating from more southern regions, such as Iraq, require less vernalization (e.g. 2–3 weeks). For a combined stratification and vernalization treatment, we typically sow the seeds and then place them at 4°C for the desired number of weeks. After approximately 3 weeks in the cold, the seeds begin to germinate. Therefore, for vernalization times

Table 16.2 Characteristics of select diploid inbred lines

Inbred line	Geographic origin	Flower under 20 h day (weeks to seed) ¹	Vernalization requirement (weeks) ²	Seed size
Bd1-1	Turkey	no	8	small
Bd2-3	Iraq	yes (12)	3	large
Bd3-1	Iraq	yes (10)	3	large
Bd18-1	Kaman Kirschir Province, Turkey (arid elev 3,000 ft)	no	8	large
Bd21 and Bd21-3	4 km from Salakudin on a road to Mosul, Iraq	yes (8)	3	large
na	northern Turkey	no	8–12	small
na	southern Turkey	no	4–6	large

¹Plants were planted in soil, placed at 4°C for 1 week to synchronize germination and then moved to a growth chamber with 20 h days.

²Number of weeks of vernalization at 4°C needed to induce flowering under short day conditions.

greater than 4 weeks we place the pots under fluorescent lighting. Vernalizing seeds/seedlings induces the plants to flower quickly while still small. Alternatively, one can vernalize larger plants if a larger amount of seed from individual plants is desired. Growth under very long day conditions (20 h light 4 h dark) overcomes the need for vernalization in a few inbred lines (Bd2-1, Bd3-1, Bd21 and Bd21-3) (Vogel et al. 2006a; Vogel and Hill 2008). Among these, the Bd21 and Bd21-3 are the most responsive and go from seed to seed in as little as 8 weeks to yield nearly six generations per year. Under these conditions the plants flower and set seed when they are approximately 15 cm tall, a size that is compatible with high density planting.

The rate of out-crossing is an important consideration when generating and maintaining large numbers of independent lines. In this regard, *Brachypodium* is a superior model. The anthers of diploid accessions rarely exert suggesting a low rate of outcrossing. This was confirmed by measuring pollen flow from transgenic to non-transgenic plants under growth chamber conditions. In a population of more than 1,000 progeny, no outcrossing was observed (unpublished). While the inbreeding nature of *Brachypodium* is an advantage for maintaining homozygous lines, it must be overcome in order to make genetic crosses. A number of researchers have crossed *Brachypodium* lines, however, the process is currently inefficient. Since grasses often shed pollen at a specific time of day, identifying when *Brachypodium* pollen is shed may increase

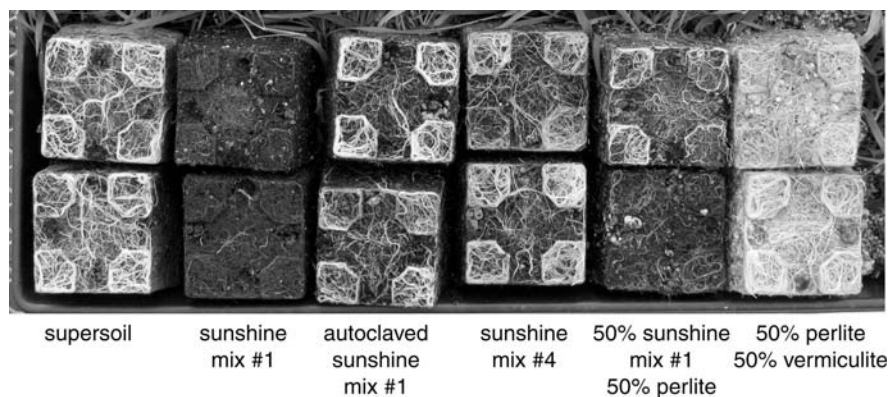


Fig. 16.2 *Brachypodium* as a model for root diseases. *Brachypodium* roots grown in different soils are shown. Plants were grown for 8 weeks and then removed from their pots and the roots photographed. Healthy white roots are evident in the plants grown in supersoil (a redwood sawdust-based mix), and 50% perlite/50% vermiculite. Plants grown in sunshine mix #1 (a peat-based mix) had severely diseased roots and *Pythium sp.* was isolated from these roots. The disease symptoms associated with sunshine mix #1 were largely eliminated by autoclaving the soil prior to planting. Incorporating perlite into the sunshine mix #1 to improve drainage also improved the symptoms. Plants grown in sunshine mix #4 (peat-based with more aggregate) were slightly diseased. Note these plants were not vernalized and thus grew much larger than plants grown for rapid seed set

crossing efficiency. This is because grass pollen is typically ephemeral and, since *Brachypodium* is normally inbreeding, *Brachypodium* pollen may be especially short lived.

Although *Brachypodium* grows well in a number of different soil types, plants that are watered excessively or left in standing water quickly develop disease symptoms. *Brachypodium* is highly susceptible to *Pythium* root rot and we have observed severe disease symptoms in some brands of commercial potting mixes (Fig. 16.2). Thus, *Brachypodium* growers may well be advised to test a few soil formulations before selecting one to grow large numbers of plants.

16.3.2 Germplasm Resources and Natural Diversity

A number of germplasm collections and several inbred lines are available to researchers. The USDA National Plant Germplasm System (NPGS) has ~30 accessions that are freely available (www.ars-grin.gov/npgs/). These collections were made many years ago and are population samples rather than inbred lines. Detailed passport data about these accessions can be found at (www.brachypodium.org). To increase the utility of this collection, inbred lines (designated by the prefix Bd) were created from five diploid and 23 polyploid NPGS accessions and have been widely distributed as a community resource (Table 16.2; Vogel et al. 2006a). A second collection assembled from many locations is maintained at the University of Aberystwyth (www.aber.ac.uk/plantpathol/germplasm.htm), and the accessions are designated by the prefix ABR. This collection contains unique material as well as re-named material from NPGS and is available through a material transfer agreement, the conditions of which may not be agreeable to some institutions. A cross-referenced list of ABR lines derived from NPGS accessions is available at (www.brachypodium.org/stocks).

Most of the material available in existing collections represents the $1n = 15$ allotetraploid cytotype. The $1n = 10$ cytotype is represented by only one accession suggesting that it may have a more restricted range, and the most useful diploid ($1n = 5$ cytotype) is represented by a handful of collections of which only five are freely available. Given that few $1n = 5$ cytotype diploids are available, a strong need exists for additional collections of diploid accessions. To begin to address this need, researchers have gathered >200 collections from many locations in Turkey (H. Budak and M. Tuna personal communication). Fortunately, much of this new Turkish material is diploid, and preliminary SSR marker analyses of these collections and existing inbred diploid lines (Bd1-1, Bd2-3, Bd3-1, Bd18-1, Bd21, Bd21-3) show a tremendous diversity both between and within wild *Brachypodium* populations (unpublished). Variability for several interesting morphological traits also has been observed among these diploid accessions including: vernalization requirements, flower initiation by

long days, size, seed pubescence, and inflorescence architecture (Fig. 16.1). The level of molecular and phenotypic diversity observed in *Brachypodium* collections indicate that *Brachypodium* can be used to identify genes responsible for natural variation in economically important traits. Furthermore, the level of polymorphism observed in SSR markers indicates that the generation of molecular markers will not be a limitation. Inbred lines have been generated from the new Turkish material, and these will be freely distributed as soon as sufficient seed is obtained.

16.3.3 Chemical and Radiation Mutagenesis

Forward and reverse genetic approaches both require large populations of mutagenized plants, and a number of protocols have been developed for using chemicals and radiation to efficiently generate large populations of mutants. Ethyl methanesulfonate (EMS) is an efficient mutagen that introduces single base changes and has been used widely to mutagenize plants. A thorough review of mutations in 192 *Arabidopsis* genes confirmed the random nature of EMS mutations and estimated the frequency at 1 mutation/170 kb of genomic DNA (Greene et al. 2003). Because EMS introduces single base changes, it can result in partial loss of function alleles that may be particularly useful when studying essential genes. We have mutagenized *Brachypodium* with EMS by adapting a method used to create a population of barley mutants that were used for TILLING (Caldwell et al. 2004). Early results indicate that EMS is an efficient mutagen for *Brachypodium* (unpublished).

Fast neutron radiation (FNR) is a complementary mutagen to EMS that introduces short deletions. Due to the larger disruptions generated by FNR, a gene disrupted by an FNR mutation may be rapidly cloned using a genome tiling array. However, since FNR typically deletes several genes, it does not lend itself to the identification of essential genes or genes located adjacent to essential genes. Early experiments indicate that *Brachypodium* can be efficiently mutagenized by FNR (D. Laudencia-Chinguanco and M. Byrne personal communication). Thus, there are no limitations in applying common mutagens to *Brachypodium*.

16.3.4 Transformation and T-DNA Tagging

Efficient transformation is a keystone of any modern model system and the extremely efficient floral dip method of *Arabidopsis* transformation is a key reason behind the tremendous success of this dicot model. Unfortunately, *Arabidopsis* is unusual in its ease of transformation and the methods used for grass transformation are much more laborious, involve extensive tissue culture and are often inefficient. Rice has set the benchmark for efficient grass

transformation and can be transformed by both *Agrobacterium* and biolistic methods at high efficiencies. A tremendous amount of work by many groups over several years increased rice transformation efficiencies from <1% in the first reports to the >40% efficiencies commonly achieved today (Tyagi and Mohanty 2000). For the other cereals (e.g. wheat and barley), however, transformation remains very inefficient. Fortunately, *Brachypodium* has proven to be very responsive to in vitro culture and current transformation efficiencies are on par with rice.

Embryogenic callus is a preferred target for transformation due to its highly regenerable nature. Thus, the development of a method for the induction of embryogenic callus from *Brachypodium* seeds and the regeneration of fertile plants from embryogenic callus was a major step toward developing *Brachypodium* transformation (Babla et al. 1995). In this study, the optimal callus-inducing medium contained LS salts, 3% sucrose and 2.5 mg l^{-1} 2,4-D. Three diploid accessions (B200, B373, B377) were found to produce embryogenic callus along with several other types of callus when mature seeds were incubated on callus-inducing media. Regeneration was observed on several common media indicating that *Brachypodium* had no unusual requirements for regeneration.

Both particle bombardment and *Agrobacterium tumefaciens* have been used to transform *Brachypodium* and each offers unique advantages and disadvantages. Particle bombardment is not dependent upon the biological limitations of *Agrobacterium*. The primary determinant of successful transformation by bombardment is the efficiency of the regeneration of plants from the bombarded explant. Thus, the transformation of embryogenic *Brachypodium* callus by particle bombardment was the next logical step in developing *Brachypodium* transformation. In the first published *Brachypodium* transformation, a polyploid accession (ABR100) was transformed via particle bombardment with an average efficiency of five transformations per g of starting embryogenic callus (Draper et al. 2001). This demonstrated that *Brachypodium* could be transformed at reasonable frequency, but raised the question of whether a diploid accession could be transformed. A more detailed account of biolistic transformation answered that question (Christiansen et al. 2005). In this paper the authors successfully transformed one diploid accession (BDR018) with an average efficiency of 5.3% of bombarded calli producing transgenic plants. They unsuccessfully attempted to transform a second diploid accession (BDR001) indicating that, similar to other plants, genotype has a substantial effect on transformation efficiency. The efficiency of this early transformation method compares favorably with the first reports of biolistic rice transformation that had an average efficiency of 3.75% (Christou et al. 1991). A serious disadvantage of biolistic transformation is the complexity of the resultant transgenic loci. Typically, these loci contain multiple copies of the inserted DNA including truncated pieces of the target DNA interspersed with genomic DNA (Kohli et al. 2003; Svitashvili and Somers 2002). These complex biolistic insertions often contain many repeats of inserted DNA and can span several

megabases of host DNA (Svitashev and Somers 2002). Such complex insertions interfere with downstream application that require relatively simple insertions (e.g. cloning flanking DNA or promoter tagging) and may lead to silencing of transgenes in later generations. Attempts to minimize the complexity of biolistic loci by using linear DNA instead of circular plasmid DNA have produced mixed results (Fu et al. 2000; Loc et al. 2002).

Agrobacterium-mediated transformation typically results in much simpler insertion patterns than biolistic transformation (for a direct comparison of methods see Dai et al. 2001; Travella et al. 2005). *Agrobacterium*-mediated transformation of both rice and *Arabidopsis* has been shown to produce low copy number transgenics with an average of ~ 1.5 insertions per line (Feldmann 1991; Jeon et al. 2000). However, the host limitations of *Agrobacterium* add to the difficulty of establishing an efficient *Agrobacterium*-mediated transformation system. Fortunately, *Brachypodium* has proven amenable to *Agrobacterium*-mediated transformation and the first report of *Agrobacterium*-mediated transformation was published in 2006 (Vogel et al. 2006a). In this study, 16 polyploid accessions and three diploid accessions were evaluated for transformability. The highest transformation efficiency (14% of the callus pieces co-cultivated with *Agrobacterium* produced transgenic plants) was achieved with the polyploid line Bd17-2. A diploid accession, PI 254867, was transformed at a much lower efficiency, 2.5%.

At the end of 2007, three papers were published online that reported very high efficiency transformation of three different *Brachypodium* lines. Two papers used inbred lines, Bd21-3 (Vogel and Hill 2008) and Bd21 (Vain et al. 2008), that were derived from the same initial USDA accession, PI 254867. Bd21-3 was selected for transformability from accession PI 254867 which presumably represents multiple individuals collected at the same location. The methods described in these two papers share a number of important similarities: media types, *Agrobacterium* strains, use of immature embryos as initial explants and they both subculture the callus several times before transformation so that each dissected embryo gives rise to many transgenic plants. This is important because of the labor involved in dissecting out immature embryos. Differences between the methods lie in the following: the use of desiccating conditions to improve transformation of Bd21-3; the formation of a yellow embryogenic callus in Bd21-3 that allows selection of the proper callus type without the aid of a microscope; the use of very small embryos and copper sulfate to improve the quality of Bd21 callus; the use of visual selection of GFP and sub-culturing callus under a microscope to improve efficiency of Bd21 selection. Average transformation efficiencies achieved (expressed as percentage of calli co-cultivated with *Agrobacterium* that produced fertile transgenic plants) were 37% for Bd21-3 and 17% for Bd21. The third paper reports extremely high average transformation efficiency, 55%, of accession BDR018 (Păcurar et al. 2008). This remarkable achievement was obtained by placing immature embryos on callus inducing media for 17 days and then co-cultivating those embryos with *Agrobacterium*. The calculated efficiency is the percentage

of dissected embryos that form fertile transgenic plants. It is significant that the embryogenic callus is not subcultured and therefore no more than one transgenic plant can arise from each dissected embryo. This increases the labor involved in generating transgenic plants when compared to the methods for Bd21-3 and Bd21 transformation. Although continued improvements in transformation efficiency, including the identification of superior genotypes, will doubtless be made, the near simultaneous publication of three high efficiency *Agrobacterium*-mediated transformation methods signals the maturation of Brachypodium transformation technology.

16.3.5 *Related Species*

The relatively small genus *Brachypodium* is estimated to have diverged from sister tribes *Triticeae* and *Poeae* 35–40 mya, and this clade has been assigned to its own tribe, *Brachypodieae*, within the subfamily *Pooideae*. Although most of the 12–15 described *Brachypodium* species have been collected from Mediterranean, European, and Eurasian locations, representatives of this genus are distributed across the globe (Catalán et al. 1995; Catalán and Olmstead 2000). Species originating in the Mediterranean include *B. distachyon* as well as the *B. retusum* and *B. phoenicoides*. In addition, one species from southern Spain, *B. boissieri*, shows substantial similarity to *B. retusum*. A single European taxon is represented by *B. rupestre*, and three species, *B. sylvaticum*, the closely related species *B. glaucovirens*, and *B. pinnatum*, are from Eurasian locations. Reports describe six additional taxa of diverse origins: *B. arbuscula* (Canary Islands), *B. kawakamii* Hayata (Taiwan), *B. mexicanum* (Mexico to Bolivia), *B. pringlei* (Central and South America), *B. bolusii* (Africa), and *B. flexum* (Africa).

All members of the *Brachypodieae* exhibit a set of common features that include: lateral stem development from the coleoptile, small chromosomes, ribosomal DNA sequence, repetitive DNA families, and shared nuclear RFLPs. However, variation in morphology, life cycle, and cytology is sufficient to clearly distinguish between species (Catalán and Olmstead 2000). Within the group, an annual life cycle is unique to *B. distachyon*. This species is also self-compatible, a trait that is shared with only two perennial species, *B. mexicanum* and *B. sylvaticum* (Khan and Stace 1999). Most of the perennial species contain long-rhizomes, however *B. mexicanum* is distinguished by being non-rhizomatous (Catalán and Olmstead 2000). Polyploidy is common among all taxa, and diploid, tetraploid, hexaploid, and octaploid species have been reported with base chromosome numbers ranging between 5 and 10 (Robertson 1981; Hasterok et al. 2004).

The phylogeny of eight *Brachypodium* species (*B. arbuscula*, *B. distachyon*, *B. mexicanum*, *B. phoenicoides*, *B. pinnatum*, *B. retusum*, *B. rupestre*, and *B. sylvaticum*) has been evaluated using several data sets including RFLP and RAPD data, chloroplast *ndhF* gene sequence, nuclear rDNA sequence, and

rDNA internal transcribed spacer (ITS) sequence (Shi et al. 1993; Hsaio et al. 1994; Catalán et al. 1995; Catalán and Olmstead 2000). Shi et al. identified an *EcoRI* site present in the rDNA of most perennial species that could be used to distinguish them from *B. distachyon* and *B. mexicanum*, but the selected markers failed to identify sufficient variation to resolve the relationship between the perennial species. An analysis using *ndhF* and ITS sequences along with RAPD data identified *B. distachyon* as the basal lineage of the group followed by the divergence of *B. mexicanum*, *B. arbuscula*, *B. retusum*, *B. rupestre*, *B. phoenicoides*, *B. pinnatum*, and then *B. sylvaticum* (Catalán and Olmstead 2000). Because of the close relationship between species in the genus *Brachypodium* researchers will be able to leverage the resources developed for *B. distachyon* to study the perennial life cycle and self-incompatibility exhibited by most other *Brachypodium* species (Khan and Stace 1999). These traits are common in the wild grasses (e.g. *Miscanthus* and switchgrass) that are being developed into biomass crops.

16.4 Genomic Resources

In order for a model system to be widely adopted, a comprehensive infrastructure of genomic resources and methods must be developed. Numerous resources have been or are currently being assembled for *Brachypodium* including: cDNA libraries, BAC libraries, a large EST collection, BAC end sequences, a high-resolution genetic linkage map, a physical map, bioinformatic resources, and most importantly, the complete genome sequence.

16.4.1 ESTs

Randomly sequencing the ends of cDNA clones to generate expressed sequence tags (ESTs) is a quick and relatively inexpensive way to learn a great deal about an unknown genome (Adams et al. 1991). Thus, it is no surprise that the first significant sequence resource for *Brachypodium* was the 20,440 ESTs deposited into Genbank in 2005 (Vogel et al. 2006b). These ESTs were derived from five cDNA libraries and represent approximately 6,000 genes. As of December 14, 2007, there were 10 grasses (including *Brachypodium*) with >20,000 ESTs in Genbank. Three of those grasses (rice, wheat, maize) had >1 million ESTs. These sequences are useful for many applications including microarrays, analysis of gene expression, and annotation of genomic sequence. The *Brachypodium* ESTs have been used to refine the phylogeny of *Brachypodium* and to identify candidates for all the genes involved in the biosynthesis of lignin monomers.

Although the initial set of *Brachypodium* ESTs are very useful, many more are required for *Brachypodium* to reach its full potential as a model system. As

part of the genome sequencing project, the U.S. Department of Energy Joint Genome Institute (JGI) is currently sequencing >180,000 additional EST sequences. To maximize the utility of this newly expanded EST collection, cDNA libraries prepared from a diverse set of tissues and treatments are being used (Todd Mockler personal communication).

16.4.2 BAC Library Resources

Bacterial artificial chromosome (BAC) libraries are useful tools for genomic analyses. Sequence data obtained from BAC libraries permit evaluation of genome content and complexity when full genome sequence is not available and can aid in assembly of genome sequencing data. Furthermore, BAC libraries are useful for comparative genomic analyses of synteny of genes in different species, and this information can be exploited to facilitate positional cloning of genes in related species. Therefore, the development of BAC library resources for *Brachypodium* has the potential to be very useful in studies of temperate cereal crops due to the close relationship between these species.

To date, six BAC libraries have been constructed for diploid *Brachypodium* accessions. The first two libraries contain a total of 9,100 clones with an average insert size of 88 kb and were derived from the genomes of accessions ABR1 (5,968 clones) and ABR5 (3,132) (Hasterok et al. 2006b). These relatively small libraries represent approximately 2 haploid genome equivalents. Two BAC libraries were constructed from inbred line Bd21, the same line that is being sequenced (Huo et al. 2006). One library was generated from HindIII digested genomic DNA and contains 36,864 clones with an average insert size of 100 kb. The other library contains 73,728 clones with an average insert size of 105 kb and was derived from BamHI digested genomic DNA. In combination, these Bd21 BAC libraries represent 29 haploid genome equivalents and provide greater than 99.99% likelihood a particular gene is included within the library. The Bd21 libraries were also used to generate BAC end sequences (BES) from 64,694 clones (average size 583 bp), and the resulting 38.2 Mbp of sequence covers ~11% of the *Brachypodium* genome (Huo et al. 2007). This sequence was used to anchor the BAC clones to the rice genome and indicated that the *Brachypodium* genome contains 45.9% GC content, approximately 18% repetitive DNA (11% with homology to known repetitive sequence and 7.3% unique to *Brachypodium*), and 21.2% coding sequence. Comparison of the BES data to *Brachypodium* and cereal crop EST databases revealed that 40% of the sequence matched ESTs with a greater number of hits within the wheat database than in the maize database. The Arizona Genomics Institute (www.genome.arizona.edu/) has constructed a library from the inbred line Bd3-1 that represents 10 genome equivalents within 36,864 clones with an average insert size of 130 kb. Two additional libraries with an average insert size of 130 kb

have been prepared from Bd21 by the Arizona Genomics Institute (M. Bevan personal communication).

An additional library exists for the perennial species *B. sylvaticum* (www.jicgenomelab.co.uk). This library contains 30,228 clones with an average insert size of 102 kb (6.6 genome equivalents, based on a genome size of 470 Mbp) (Foote et al. 2004). From this library, repetitive DNA content was estimated to be approximately 50% and analyses demonstrated that synteny was maintained between rice, wheat, and *B. sylvaticum* BAC contigs over several regions of chromosome 9. The percentage of repetitive DNA in *B. sylvaticum* is much higher than in *B. distachyon* and largely explains the larger size of the *B. sylvaticum* genome.

16.4.3 Physical and Genetic Maps

There are currently no published physical or genetic maps for *Brachypodium*. However, much progress in this direction has been made and it is anticipated that both a physical and a genetic map will be published shortly. A physical map has been constructed from two of the Bd21 BAC libraries mentioned above (Huo et al. 2006, 2007). This map contains over 50,000 BAC clones assembled into ~600 contigs (M. Luo personal communication). It is anticipated that the map will be published in early 2009 and become available at: (phymap.ucdavis.edu:8080/brachypodium/). In addition, a second physical map using two different Bd21 libraries has recently been constructed (M. Bevan personal communication). Rapid progress toward a genetic map is also being made. A large community collaboration is using ~200 markers to create the first genetic linkage map, and it is anticipated that this effort will be complete shortly (D. Garvin personal communication). As a measure of the rapid progress in *Brachypodium* research, a National Science Foundation funded project to create a high-density SNP based map is progressing rapidly even before the first generation linkage map is finished. The goal of this project is to map ~1,000 SNP markers. As of January 2008, 1,900 SNPs at 625 loci have been identified and the map is anticipated to be completed by mid 2008 (unpublished). These mapping resources will greatly aid in the final assembly and verification of the complete genome sequence and also will aid positional cloning experiments.

Linking individual BACs contained in physical contigs and ultimately genomic sequences to specific chromosomes can be accomplished through a technique called “BAC landing.” In this technique, entire BACs are fluorescently labeled and used for FISH. In this fashion, BACs were assigned to specific chromosomes, and 32 of 39 BACs hybridized to a single locus underscoring the compact nature of the *Brachypodium* genome (Hasterok et al. 2006b). A more extensive application of the technique will be highly instructive in verifying the whole genome assembly.

16.4.4 Whole Genome Sequencing

A completely sequenced genome is a requirement for a modern model system and underpins a host of tools including efficient map-based cloning, sequence indexed T-DNA populations, gene chips and reverse genetic approaches including TILLING and RNAi. Plans for the development of Brachypodium as a model to accelerate the domestication of grasses (e.g. switchgrass and *Miscanthus*) for use as biomass crops were spelled out in a U.S. Department of Energy report on the research needed to establish a domestic biofuel industry (DOE 2006). As a result, the JGI (www.jgi.doe.gov/) approved a proposal to sequence the Brachypodium genome through their Community Sequencing Program for 2007. The JGI is using a whole genome shotgun sequencing approach based upon Sanger sequencing technology with a final target of 8x genome coverage. The final assembly will incorporate all of the mapping and BAC sequence resources and promises to be of very high quality. A preliminary 4x draft sequence has been released through (www.brachypodium.org) and (www.modelcrop.org). It is anticipated that the final 8x genome assembly will be completed in 2008.

16.4.5 Bioinformatic Resources

To fully utilize the approaching avalanche of genomic data, it will be necessary to develop the appropriate bioinformatic infrastructure. A Brachypodium-specific web portal (www.brachypodium.org) that provides links to numerous sources of Brachypodium information has been established. This website also houses a newsgroup that links the Brachypodium community. As mentioned above, the 4x Brachypodium sequence is housed on two databases (www.brachybase.org and www.modelcrop.org). In addition to the 4x sequence, these databases contain or will contain tracks that place genetic markers, BAC clones, ESTs, T-DNA insertion sites, sequences from other species, and other applicable data in the context of the Brachypodium genomic sequence. Readers are directed to these sites for the most detailed resources available. Other websites that contain project specific information include the “Brachyomics” website (www.aber.ac.uk/plantpathol/brachyomics.htm) and a website describing projects underway at the USDA-ARS Genomics and Gene Discovery Unit (brachypodium.pw.usda.gov/).

16.5 Applications of Brachypodium as a Model for Grass Research

The close relationship of Brachypodium to wheat holds the promise that the simple genome of Brachypodium can be used as a roadmap to navigate the complex wheat genome. The application of Brachypodium as such a structural model is dependent upon the conservation of colinearity between wheat and

Brachypodium. Thus, since wheat is more closely related to Brachypodium than to rice, it is anticipated that Brachypodium will serve as a better structural model than rice. However, colinearity will vary from locus to locus, and the utility of Brachypodium will vary depending upon the region examined. Thus, including the rice genome in a three-way comparison with Brachypodium and wheat may be more informative in some situations.

In addition to serving as a structural model for wheat, Brachypodium can serve as a functional model for grasses in general. In this capacity, it is not necessary to have an extremely close evolutionary relationship. It is only necessary to share the traits/properties and genes under study. For example, Brachypodium possess the type 2 cell wall (Carpita 1996) typical of all grasses and thus would be a suitable model to study this facet of grass biology. By contrast, Arabidopsis contains a type 1 cell wall typical of the dicots and therefore would be a poor choice to study the unique aspects of the grass cell wall.

16.5.1 *Brachypodium* as Structural Model for Wheat and Barley Genomics

The polyploid nature, large size and highly repetitive nature of the wheat genome present extreme challenges to researchers studying specific genes or genomic regions. A simpler, yet closely related, genome could serve as a roadmap to accelerate research on the complex wheat genome by providing a frame of reference and sequences from which markers in intervals of interest can be developed. Both *B. sylvaticum* and *B. distachyon* have already been used for this purpose. With the availability of the complete Brachypodium genome sequence, Brachypodium will be increasingly employed in this capacity. The use of a simple genome as a model for more complex genomes is dependent on a high degree of colinearity. As the phylogenetic distance increases between two species, assessment of colinearity becomes increasingly difficult, and within 1 Mb of sequence, the comparison of gene order between rice and wheat becomes less reliable (Foote et al. 2004). In this regard, Brachypodium has an advantage over rice because rice and wheat diverged approximately 50 million years ago (Gaut 2002) whereas wheat and Brachypodium diverged about 35 million years ago (Bossolini et al. 2007). As predicted based on evolutionary distance, a comparison of a 371 kb genomic sequence from *B. sylvaticum* to the orthologous rice sequence and to the segregation of wheat genes in this interval revealed that the gene content and gene order of the wheat region was closer to *B. sylvaticum* than to rice (Bossolini et al. 2007). Specifically, of the 15 wheat genes found in this interval, 10 had orthologs in the *B. sylvaticum* sequence and nine had orthologs in the rice interval. The order of the shared genes was the same in *B. sylvaticum* and wheat whereas there was a large inversion in rice. This suggests that Brachypodium will be a better model of the wheat genome than is rice.

Three examples of using *B. sylvaticum* sequence to map or clone wheat or barley genes have been published. *B. sylvaticum* BACs and rice genomic sequence were used in combination to clone the *Ph1* locus from wheat (Griffiths et al. 2006). In this study the authors used markers in a region of rice to define the breakpoints of wheat deletion lines that defined a region containing the *Ph1* locus. However, they were unable to map half of the rice markers in the wheat genome due to lack of sequence conservation. To map the remaining markers, they used the rice sequences to obtain the orthologous sequences from a *B. sylvaticum* library. They were then able to map all of these *B. sylvaticum* markers in wheat to narrow the interval containing the *Ph1* locus. That *B. sylvaticum* sequences could be mapped directly onto the wheat genome where the orthologous rice sequences failed underscores the close relationship of *B. sylvaticum* to wheat. The *B. sylvaticum* genome was also used to narrow the interval containing the *Lr34/Yr18* disease resistance locus in wheat (Spielmeyer et al. 2007). In this case, one of the *B. sylvaticum* BACs sequenced by Bosollini et al. 2007 (discussed above) was found to overlap the *Lr34/Yr18* region in *B. sylvaticum* and the authors were able to use this sequence to narrow the interval containing *Lr34/Yr18*. *B. sylvaticum* BACs were also used in the cloning of the *Ppd-H1* gene from barley though in this case the *B. sylvaticum* sequence primarily confirmed results obtained with rice sequence (Turner et al. 2005). From these examples, it is apparent that Brachypodium will be very useful as a roadmap for wheat. The extent of this utility will become apparent when a larger comparison of the synteny between wheat and Brachypodium is made.

16.5.2 *Brachypodium as a Functional Model*

A systematic approach to analyses of gene function is the goal that drives the development of functional model systems. The close relationship to temperate cereals and forage grasses paired with available and emerging genomic tools make Brachypodium an attractive model for these purposes. Examples of the merit of Brachypodium as a functional model are beginning to emerge. A recent study of the floral repressor *Terminal Flower 1* demonstrated that the orthologous genes from Arabidopsis (*TFL1*) and *Lolium perenne* (*LpTFL1*) both function to delay flowering in Brachypodium (Olsen et al. 2006). The grass gene, *LpTFL1*, mediated on average a 14 day longer delay in flowering than the gene from the dicot Arabidopsis suggesting that Brachypodium is a more appropriate system in which to pose questions regarding regulatory pathways of monocots. Furthermore, the inbreeding life cycle and short generation time of Brachypodium significantly accelerated the analysis of transgenic plants. Within one year of transformation, T₁ transgenic plants were ready for analysis, a feat that could not be accomplished with *Lolium perenne* due to low transformation efficiency, the requirement for extensive vernalization, and

self-incompatibility. Another application of *Brachypodium* to study development was the isolation of *Brachypodium* orthologs of *WUSCHEL* homeobox (WOX) gene family members (Nardmann et al. 2007). The authors had previously identified the members of the WOX gene family from rice and maize and thus with the *Brachypodium* sequences they were able to construct a phylogenetic tree containing members from the three major radiations of the family Poales. This comparison revealed gene duplications common to all grasses that were not present in dicots.

Initial surveys of *Brachypodium* responses to pathogen challenge indicate that this species is also well suited to the investigation of host-pathogen interactions in grasses. In the most comprehensive study to date, twenty-one diploid *Brachypodium* accessions were screened for susceptibility to three strains of the rice blast pathogen *Magnaporthe grisea* (Routledge et al. 2004). These experiments revealed responses ranging from a highly localized hypersensitive response to full susceptibility with cytology similar to that observed in rice. One accession, ABR5, presented full resistance to the *M. grisea* strain Guy-11. Based on segregation of resistance in a cross with the susceptible accession ABR1, the resistance appears to be mediated by a single locus. Hallmarks typically used to measure disease resistance in *Arabidopsis*, such as PR protein expression, callose deposition, and the appearance of autofluorescence and granular cytoplasm in infected regions, were also evident in the *Brachypodium* response to *M. grisea*. Virulence tests for several rust pathogens (*Puccinia striiformis hordei*, *Puccinia striiformis triticae*, *Puccinia recondita hordei*, *Puccinia recondita triticae*, *Puccinia coronata*) revealed substantial variation in resistance response between different *Brachypodium* accessions (Draper et al. 2001), and susceptibility to head blight caused by *F. graminearum* (Garvin 2007) and root rot resulting from infection by *Pythium* species also have been observed (Fig. 16.2).

In contrast to interactions between *Brachypodium* and *M. grisea*, the powdery mildew pathogen *Blumeria graminis* failed to elicit any observable disease symptoms on *Brachypodium* accessions (Draper et al. 2001). This suggests a non-host rather than race-specific resistance to this pathogen. Robust, non-host resistance it is thought to be mediated by multiple genes that together control a wide range of functions including the production of toxic agents or the absence of metabolites or signaling molecules required by the pathogen. Therefore, inactivation of multiple components would be required to render a plant susceptible to infection. Study of non-host resistance to *B. graminis* is being pursued in *Arabidopsis* as an alternative to the introduction of short-lived, race-specific disease resistance genes to crop species (Collins et al. 2003). These results suggest that *Brachypodium* can serve as a model to initiate similar studies in grasses.

Despite the agronomic importance of cereal viruses, information regarding interactions of these pathogens with *Brachypodium* is conspicuously lacking. One undefined, spherical virus has been described in *B. sylvaticum* (Edwards et al. 1985), and initial studies in *Brachypodium* have identified accessions that

are either susceptible or resistant to challenges with *Barley stripe mosaic hordeivirus* (A. Jackson, personal communication).

In addition to serving as a model for plant-pathogen interactions, Brachypodium can serve as a model for the interaction between plants and mycorrhizal fungi. Since mycorrhizal interactions can greatly improve plant productivity and decrease costs associated with the application of fertilizer, this is a very important area for plants in general and for the use of grasses as biomass crops in particular. Despite its importance, this area has been understudied in part because of the difficulties of working with roots and in part because *Arabidopsis* does not form mycorrhizal associations. Since Brachypodium forms robust mycorrhizal interactions (M. Harrison personal communication) it will no doubt serve as a powerful model these symbiotic relationships.

Brachypodium is also being used as a model for responses to wounding and insect attack. A proteinase inhibitor, *Bdpin1*, was identified from a cDNA library prepared from wounded Brachypodium leaves (Mur et al. 2004). Local and systemic expression of *Bdpin1* in response to wounding demonstrated the presence of long-distance signaling in grasses. *Bdpin1* was also induced in response to methyl jasmonate and *M. grisea*, but not in response to the salicylic acid analog benzothiadiazole. Taken together, these results indicate that Brachypodium can serve as a model for grass development and the interactions of grasses with pathogens, insects and symbionts.

16.6 Future Prospects and Directions

The future of Brachypodium as a model system is very bright. The development of a base suite of tools (e.g. complete genome sequence, inbred lines, BAC and cDNA libraries, facile transformation methods, physical and genetic maps) that allow researchers to rapidly utilize Brachypodium to study their questions of interest has been extraordinarily rapid. This is due to a favorable alignment of several factors including: an increased interest in grasses as feedstocks for bioenergy; the need for a tractable grass model as a surrogate for the large, difficult to handle grasses proposed as biomass crops; the recognition that the Brachypodium genome could be used as a comparative tool to probe the extremely large and highly repetitive wheat genome; the availability of low cost high-throughput sequencing capacity through the Community Sequencing Program; and the willingness of a handful of researchers to invest in developing freely-available community resources. The momentum of these initial developments has carried over into the development of next generation resources that are either currently in early stages of development (e.g. insertional mutants, TILLING resources, and elucidation of small RNAs and microRNAs) or being planned and proposed to funding agencies (e.g. microarrays, resequencing of additional lines). This rapid progress is likely to continue for the foreseeable future as the number of researchers using Brachypodium continues to grow exponentially.

Acknowledgments This work was supported by USDA CRIS project 5325-21000-013-00 “Biotechnological Enhancement of Energy Crops” and by the Office of Science (BER), U.S. Department of Energy, Interagency Agreement No. DE-AI02-07ER64452.

References

- Adams, M., Kelley, J., Dubnick, M., Polymeropoulos, M., Xiao, H., Merrill, C., Wu, A., Olde, B., Moreno, R., Kerlavage, a., McCombie, W. and Venter, J. (1991) Complementary DNA sequencing: expressed sequence tags and human genome project. *Science* 252, 1651–1656.
- Bablak, P., Draper, J., Davey, M.R. and Lynch, P.T. (1995) Plant regeneration and micro-propagation of *Brachypodium distachyon*. *Plant Cell, Tiss. Org. Cult.* 42, 97–107.
- Bennett, M.D. and Leitch, I.J. (2005) Nuclear DNA amounts in Angiosperms: progress, problems and prospects. *Ann. Botany* 95, 45–90.
- Bossolini, E., Wicker, T., Knobel, P.A. and Keller, B. (2007) Comparison of orthologous loci from small grass genomes *Brachypodium* and rice: implications for wheat genomics and grass genome annotation. *Plant J.* 49, 704–717.
- Caldwell, D.G., McCallum, N., Shaw, P., Muehlbauer, G.J., Marshall, D.F. and Waugh, R. (2004) A structured mutant population for forward and reverse genetics in Barley (*Hordeum vulgare* L.). *Plant J.* 40, 143–150.
- Carpita, N.C. (1996) Structure and biogenesis of the cell walls of grasses. *Ann. Rev. Plant Phys. Plant Mol. Biol.* 47, 445–476.
- Catalán, P. and Olmstead, R.G. (2000) Phylogenetic reconstruction of the genus *Brachypodium* P. Beauv. (Poaceae) from combined sequences of chloroplast *ndhF* gene and nuclear ITS. *Plant Syst. Evol.* 220, 1–19.
- Catalán, P., Ying, S., Armstrong, L., Draper, J. and Stace, C.A. (1995) Molecular phylogeny of the grass genus *Brachypodium* P. Beauv. based on RFLP and RAPD analysis. *Botanical J. Linn. Soc.* 117, 263–280.
- Christiansen, P., Didion, T., Andersen, C.H., Folling, M. and Nielsen, K.K. (2005) A rapid and efficient transformation protocol for the grass *Brachypodium distachyon*. *Plant Cell Rep.* 23, 751–758.
- Christou, P., Ford, T.L. and Kofron, M. (1991) Production of transgenic rice (*Oryza sativa* L.) plants from agronomically important indica and japonica varieties via electric discharge particle acceleration of exogenous DNA into immature zygotic embryos. *Bio. Technol.* 9, 957–962.
- Collins, N.C., Thordal-Christensen, H., Lipka, V., Bau, S., Kombrink, E., Qiu, J-L., Hückelhoven, R., Stein, M., Freialdenhoven, A., Somerville, S.C. and Schulze-Lefert, P. (2003) SNARE-protein-mediated disease resistance at the plant cell wall. *Nature* 425, 973–977.
- Dai, S., Zheng, P., Marmey, P., Zhang, S., Tian, W., Chen, S., Beachy, R.N. and Fauquet, C. (2001) Comparative analysis of transgenic rice plants obtained by *Agrobacterium*-mediated transformation and particle bombardment. *Mol. Breed.* 7, 25–33.
- DOE. (Ed.). (2006). *Breaking the Biological Barriers to Cellulosic Ethanol: A Joint Research Agenda*, U.S. Department of Energy, Office of Science and Office of Energy Efficiency.
- Döring, E., Schneider, J., Hilu, K.W. and Röser, M. (2007) Phylogenetic relationships in the Aveneae/Poeae complex (Pooideae, Poaceae). *Kew Bull.* 62, 407–424.
- Draper, J., Mur, L.A.J., Jenkins, G., Ghosh-Biswas, G.C., Bablak, P., Hasterok, R. and Routledge, A.P.M. (2001) *Brachypodium distachyon*. A new model system for functional genomics in grasses. *Plant Phys.* 127, 1539–1555.
- Edwards, M., Cooper, J., Massalski, P. and Green, B. (1985) Some properties of a virus-like agent found in *Brachypodium sylvaticum* in the United Kingdom. *Plant Pathol.* 34, 95–104.

- Feldmann, K.A. (1991) T-DNA insertion mutagenesis in *Arabidopsis*: mutational spectrum. *Plant J.* 1, 71–82.
- Foote, T.N., Griffiths, S., Allouis, S. and Moore, G. (2004) Construction and analysis of a BAC library in the grass *Brachypodium sylvaticum*: its use as a tool to bridge the gap between rice and wheat in elucidating gene content. *Funct. Integr. Genomics* 4, 26–33.
- Fu, X., Duc, L.T., Fontana, S., Bong, B.B., Tinjuangjun, P., Sudhakar, D., Twyman, R.M., Christou, P. and Kohli, A. (2000) Linear transgene constructs lacking vector backbone sequences generate low-copy-number transgenic plants with simple integration patterns. *Transgenic Res.* 9, 11–19.
- Garvin, D.F. (2007) *Brachypodium*: a new monocot model plant system emerges. *J. Sci. Food Agric.* 87, 1177–1179.
- Gaut, B.S. (2002) Evolutionary dynamics of grass genomes. *New Phytol.* 154, 15–28.
- Greene, E.A., Codomo, C.A., Taylor, N.E., Henikoff, J.G., Till, B.J., Reynolds, S.H., Enns, L.C., Burtner, C., Johnson, J.E., Odden, A.R., Comai, L. and Henikoff, S. (2003) Spectrum of chemically induced mutations from a large-scale reverse-genetic screen in *Arabidopsis*. *Genetics* 164, 731–740.
- Griffiths, S., Sharp, R., Foote, T.N., Bertin, I., Wanous, M., Reader, S., Colas, I. and Moore, G. (2006) Molecular characterization of *Ph1* as a major chromosome pairing locus in polyploid wheat. *Nature* 439, 749–752.
- Hasterok, R., Draper, J. and Jenkins, G. (2004) Laying the cytotoxic foundations of a new model grass, *Brachypodium distachyon* (L.) beauv. *Chromosome Res.* 12, 397–403.
- Hasterok, R., Dulawa, J., Jenkins, G., Leggett, M. and Langdon, T. (2006a) Multi-substrate chromosome preparations for high throughput comparative FISH. *BMC Biotechnol.* 6, 20.
- Hasterok, R., Marasek, A., Donnison, I.S., Armstead, I., Thomas, A., King, I.P., Wolny, E., Idziak, D., Draper, J. and Jenkins, G. (2006b) Alignment of the genomes of *Brachypodium distachyon* and temperate cereals and grasses using bacterial artificial chromosome landing with fluorescence *in situ* hybridization. *Genetics* 173, 349–362.
- Hsiao, C., Chatterton, N.J., Asay, K.H. and Jensen, K.B. (1994) Phylogenetic relationships of 10 grass species: an assessment of phylogenetic utility of the internal transcribed spacer region in nuclear ribosomal DNA in monocots. *Genome* 37, 112–120.
- Huo, N., Gu, Y., Lazo, G., Vogel, J., Coleman-Derr, D., Luo, M., Thilmony, R., Garvin, D. and Anderson, O. (2006) Construction and characterization of two BAC libraries from *Brachypodium distachyon*, a new model for grass genomics. *Genome* 49, 1099–1108.
- Huo, N., Lazo, G.R., Vogel, J.P., You, F.M., Ma, Y., Hayden, D.M., Coleman-Derr, D., Hill, T.A., Dvorak, J., Anderson, O.D., Luo, M. and Gu, Y.Q. (2007) The nuclear genome of *Brachypodium distachyon*: analysis of BAC end sequences. *Funct. Integr. Genomics* 8, 135–147.
- Jeon, J., Lee, S., Jung, K., Jun, S., Jeong, D., Lee, J., Kim, C., Jang, S., Lee, S., Yang, K., Nam, J., An, K., Han, M., Sung, R., Choi, H., Yu, J., Choi, J., Cho, S., Cha, S., Kim, S. and An, G. (2000) T-DNA insertional mutagenesis for functional genomics in rice. *Plant J.* 22, 561–570.
- Kellogg, E.A. (1998) Relationships of cereal crops and other grasses. *Proc. Natl. Acad. Sci. USA* 95, 2005–2010.
- Kellogg, E.A. (2001) Evolutionary history of the grasses. *Plant Phys.* 125, 1198–1205.
- Khan, M.A. and Stace, C.A. (1999) Breeding relationships in the genus *Brachypodium* (Poaceae: Pooideae). *Nordic J. Bot.* 19, 257–269.
- Kohli, A., Twyman, R.M., Abranches, R., Wegel, E., Stoger, E. and Christou, P. (2003) Transgene integration, organization and interaction in plants. *Plant Mol. Biol.* 52, 247–258.
- Loc, N.T., Tinjuangjun, P., Gatehouse, A.M.R., Christou, P. and Gatehouse, J.A. (2002) Linear transgene constructs lacking vector backbone sequences generate transgenic rice plants which accumulate higher levels of proteins conferring insect resistance. *Mol. Breed.* 9, 231–244.

- Mur, L.A.J., Xu, R., Casson, S.A., Stoddart, W.M., Routledge, A.P.M. and Draper, J. (2004) Characterization of a proteinase inhibitor from *Brachypodium distachyon* suggests the conservation of defence signalling pathways between dicotyledonous plants and grasses. *Mol. Plant Pathol.* 5, 267–280.
- Nardmann, J., Zimmermann, R., Durantini, D., Kranz, E. and Werr, W. (2007) WOX gene phylogeny in poaceae: a comparative approach addressing leaf and embryo development. *Mol. Biol. Evol.* 24, 2474–2484.
- Olsen, P., Lenk, I., Jensen, C.S., Petersen, K., Andersen, C.H., Didion, T. and Nielsen, K.K. (2006) Analysis of two heterologous flowering genes in *Brachypodium distachyon* demonstrates its potential as a grass model plant. *Plant Sci.* 170, 1020–1025.
- Păcurar, D.I., Thordal-Christensen, H., Nielsen, K.K. and Lenk, I. (2008) A high-throughput *Agrobacterium*-mediated transformation system for the grass model species *Brachypodium distachyon* L. *Transgenic Res.* 17, 965–975.
- Robertson, I.H. (1981) Chromosome numbers in *Brachypodium* Beauv. (Gramineae). *Genetica* 56, 55–60.
- Rokas, A., Williams, B.I., King, N. and Carroll, S.B. (2003) Genome-scale approaches to resolving incongruence in molecular phylogenies. *Nature* 425, 798–804.
- Routledge, A.P.M., Shelley, G., Smith, J.V., Draper, J., Mur, L.A.J. and Talbot, N.J. (2004) *Magnaporthe grisea* interactions with the model grass *Brachypodium distachyon* closely resemble those with rice (*Oryza sativa*). *Mol. Plant Pathol.* 5, 253–265.
- Shi, Y., Draper, J. and Stace, C. (1993) Ribosomal DNA variation and its phylogenetic implication in the genus *Brachypodium* (Poaceae). *Plant Syst. Evol.* 188, 125–138.
- Spielmeier, W., Singh, R.P., McFadden, H., Wellings, C.R., Huerta-Espino, J., Kong, X., Appels, R. and Lagudah, E.S. (2007) Fine scale genetic and physical mapping using interstitial deletion mutants of *Lr34/Yr18*: a disease resistance locus effective against multiple pathogens in wheat. *Theor. Appl. Genet.* 116, 481–490.
- Svitashev, S.K. and Somers, D.A. (2002) Characterization of transgene loci in plants using FISH: a picture is worth a thousand words. *Plant Cell Tiss. Org. Cult.* 69, 205–214.
- Travella, S., Ross, S.M., Harden, J., Everett, C., Snape, J.W. and Harwood, W.A. (2005) A comparison of transgenic barley lines produced by particle bombardment and *Agrobacterium*-mediated techniques. *Plant Cell Rep.* 23, 780–789.
- Turner, A., Beales, J., Faure, S., Dunford, R.P. and Laurie, D.A. (2005) Botany: the pseudo-response regulator Ppd-H1 provides adaptation to photoperiod in barley. *Science* 310, 1031–1034.
- Tyagi, A.K. and Mohanty, A. (2000) Rice transformation for crop improvement and functional genomics. *Plant Sci.* 158, 1–18.
- Vain, P., Worland, B., Thole, V., McKenzie, N., Alves, S., Opanowicz, M., Fish, L., Bevan, M. and Snape, J. (2008) *Agrobacterium*-mediated transformation of the temperate grass *Brachypodium distachyon* (genotype Bd21) for T-DNA insertional mutagenesis. *Plant Biotechnol. J.* 6, 236–245.
- Vogel, J.P., Garvin, D.F., Leong, O.M. and Hayden, D.M. (2006a) *Agrobacterium*-mediated transformation and inbred line development in the model grass *Brachypodium distachyon*. *Plant Cell Tiss. Org. Cult.* 85, 199–211.
- Vogel, J.P., Gu, Y., Twigg, P., Lazo, G., Laudencia-Chingcuanco, D., Hayden, D., Donze, T., Vivian, L., Stamova, B. and Coleman-Derr, D. (2006b) EST sequencing and phylogenetic analysis of the model grass *Brachypodium distachyon*. *Theor. Appl. Genet.* 113, 186–195.
- Vogel, J. and Hill, T. (2008) High-efficiency *Agrobacterium*-mediated transformation of *Brachypodium distachyon* inbred line Bd21-3. *Plant Cell Rep.* 27, 471–478.

Chapter 17

Comparative Genomics in the Triticeae

Catherine Feuillet and Jérôme Salse

Abstract The genomes of grasses are very different in terms of size, ploidy level and chromosome number. Among them, the Triticeae species (wheat, barley, rye) have some of the largest and complex genomes. Comparative mapping studies between rice, maize, sorghum, barley and wheat have pioneered the field of plant comparative genomics a decade ago. They showed that the linear order (colinearity) of genetic markers and genes is very well conserved opening the way to accelerated map-based cloning and defining rice as a model for grasses. More recently, the availability of BAC libraries and large sets of genomic sequences including the completion of the rice genome have permitted micro-colinearity studies that revealed rearrangements between the grass genomes and provided some insights into mechanisms that have shaped their genome during evolution. This review summarizes a decade of comparative genomics studies in grasses with a special emphasis on the wheat and barley genomes.

17.1 Introduction

Cereals (grass species that are cultivated for their edible seeds) such as wheat, rice, maize, barley, oat, sorghum or millets constitute over 50% of the total crop production worldwide (<http://www.fao.org/>), and their grains have represented since the beginning of agriculture one of the most important renewable resources for human food and domestic animal feed in most of the civilizations. Moreover, cereal seeds and straw represent a potential source of non-food products and for bioenergy production to supplement or replace fossil energies in the future. All cereal crop species are members of the grass (Poaceae) family that is the fourth largest family of flowering plants. With about 10,000 species growing under nearly all climates and latitudes, grasses exceed all other families

J. Salse (✉)

INRA-UBP UMR 1095, Domaine de Crouelle, 234 Avenue du Brezet, F-63100
Clermont-Ferrand, Genetics, Diversity and Ecophysiology of Cereals, France
e-mail: jsalse@clermont.inra.fr

in ecological dominance and economic importance. Within the grass family, the cereals are represented in four of the five main sub-families (Fig. 17.1): Sorghum, maize, pearl millet and foxtail millet are members of the Panicoideae; finger millet belongs to the Chloridoideae; rice to the Ehrhartoideae while wheat, barley, oat and rye are Pooideae representatives. Fossil data and phylogenetic studies, have estimated that the grasses have diverged from a common ancestor 50–70 MYA (for reviews see Kellogg 2001; Gaut 2002).

Among the Pooideae subfamily, the Triticeae tribe comprises some of the most important crops worldwide e.g. wheat (*Triticum aestivum* L.) barley (*Hordeum vulgare* L.) and rye (*Secale cereale* L.). They are characterized by large genome sizes (5,500 Mb for barley, 8,000 Mb for rye and 17,000 Mb for bread wheat; Bennett and Smith 1976) that are composed of more than 80% of repetitive DNA (Smith and Flavell 1975). It is estimated that barley and wheat have diverged 10–14 MYA while wheat and rye derived from a common ancestor less than 6 MYA (Kellogg 2001, Fig. 17.1).

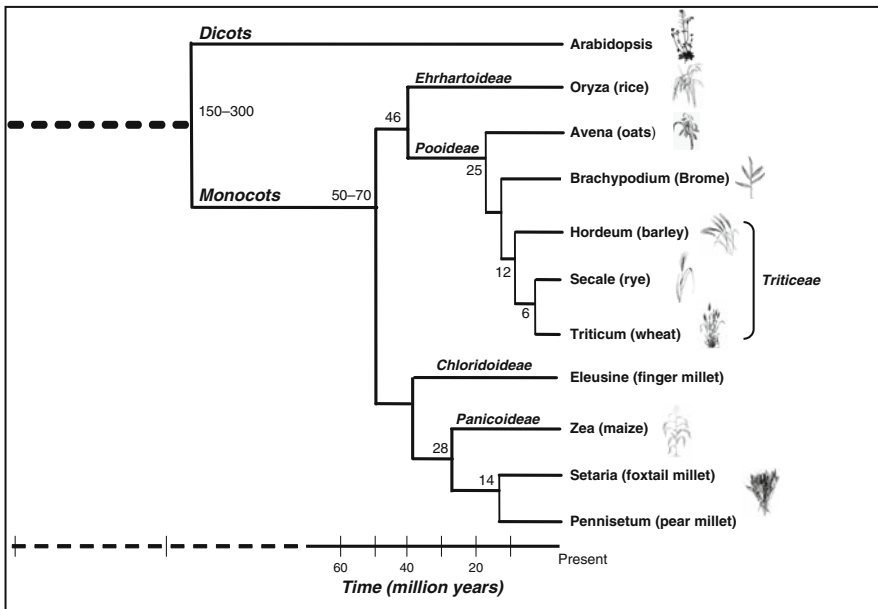


Fig. 17.1 Phylogenetic relationships between grasses and Triticeae species. Divergence times from a common ancestor are indicated on the branches of the phylogenetic tree (in millions years)

Comparative genomics, i.e. studies of the relationships between genomes of different species, permits to identify the portion of genomes that are conserved and those that are unique allowing to relate specific changes in genome structure and content to differences in the biology of the different species. It provides insight into the mechanisms of genome evolution and speciation as well as tools

for a variety of studies and applications ranging from the densification of DNA markers on genetic maps to the identification of conserved genes and regulatory sequences. Comparative genomics in families that have a relatively recent history such as the Triticeae have a great potential because they allow to access and understand the basis of diversity and adaptation for a better exploitation of the genetic resources in crop improvement. Comparative analyses between the Triticeae species or between the Triticeae and other cereals have been the focus of intense research in the past decade and have pioneered this field of research. The first results have indicated a good conservation of the markers order at the genetic map level (macrocolinearity) and have promoted rice as a reference genome for the Triticeae and more generally for grasses. The release of the first rice genome sequence drafts in 2002 and the development of a number of genomic resources (EST collections, BAC libraries) from the Triticeae have then allowed further inter- and intra-specific comparative studies at the sequence level both at the genome or chromosomes scale and at target loci. These studies shed new light into the degree of conservation between the Triticeae and the other grass genomes and provided the first insights into the mechanisms that have shaped these genomes within 50–70 million years of evolution. It also contributed to the map-based cloning of the first genes in wheat and barley by using markers from other grass species and rice in particular, to increase marker density and reduce the genetic intervals around the target genes.

A number of reviews have already been published on comparative genomics in the cereals (Paterson et al. 2005; Salse and Feuillet 2007) and this chapter will review more specifically comparative genomics studies performed in the past decade and involving at least one of the Triticeae genomes. It will discuss what was learned from these studies about Triticeae and grass genomes organisation and evolution, describe how this knowledge can be applied to support gene discovery and Triticeae crop improvement and finally what can be expected in the next years in terms of comparative studies with the sequencing of the Triticeae genomes.

17.2 Comparative Genomics at the Genome Scale:

Macrocolinearity

Early comparative genetic mapping studies have indicated that despite large differences in ploidy level, chromosome number, and haploid DNA content, the linear order (colinearity) of markers remained largely conserved between grass species including the Triticeae over several millions of years of evolution (reviewed in Devos and Gale 2000; Feuillet and Keller 2002; Salse and Feuillet 2007). The estimated level of colinearity has evolved over the years at the same time that the level of resolution of the analysis has increased with the saturation of genetic maps with new markers and the availability of rice genome sequences after 2002.

17.2.1 Marker Based Macrocolinearity Studies

Initial comparisons between the genomes of all important grass species were performed originally with restriction fragment length polymorphisms (RFLP) markers. This provided compelling evidence that except for few large rearrangements, the linear order of markers was conserved on the chromosomes (macrocolinearity) despite 50–70 million years of divergent evolution. Comparative RFLP mapping reported a good level of colinearity between Triticeae species at the genome level and a reference Triticeae consensus map has been developed from common RFLP markers genetically mapped in -diploid, tetraploid and hexaploid- wheat, barley and rye (Van Deynze et al. 1995; Devos et al. 1993b; Herrmann et al. 1996; Hohmann et al. 1994; Dubcovsky et al. 1996). These RFLP studies led to the establishment of orthologous relationships between the three Triticeae genomes and to the adoption of a common nomenclature for the seven basic chromosomes in these species. It also contributed to the characterization of translocations involving chromosome arms 4AL, 5AL and 7BS in wheat which were found identified in rye on chromosomes 4RL, 5RL and 7RS (Naranjo et al. 1987; Koller and Zeller 1976; Naranjo and Fernandez-Rueda 1991; Liu et al. 1992; Devos et al. 1993a,b). These data were integrated into the famous “crop circles” (Moore et al. 1995; Devos and Gale 1997) that provided a representation of the relationships between orthologous chromosomes in eight species belonging to three grass subfamilies: rice (Ehrhartoideae), foxtail millet, sugar cane, sorghum, pearl millet, maize (Panicoideae) as well as the Triticeae (considered as a single genetic system with seven chromosomes) and oat (Pooideae).

Macrocolinearity studies indicated a generally good level of conservation between the grass genomes and led to consider them as a single genetic system built from 30 rice linkage blocks that possibly represented linkage blocks of the ancestral grass genome (Moore et al. 1995). However, these results were obtained from low resolution genetic maps with an average of one marker every 10 cM that only allowed the detection of dramatic rearrangements. Moreover, the maps were constructed with low copy RFLP markers that were selected for their ability to provide a signal in cross hybridizations limiting the detection of whole or partial genome duplication events and making it difficult to assess orthologous and paralogous relationships of gene families. Finally, as comparisons based on the genetic maps overemphasis polymorphic regions, the overall genomes were not evenly represented especially for the centromeric regions.

A reassessment of the colinearity among the grass genomes was performed by Gaut in 2002 who collated data from different comparative studies to estimate the probability for one marker found in the vicinity of another to be in a colinear region. The results indicated that the average probability was not very high between the grass families with about 50% on average between the Triticeae and rice. It increased within families, between more closely related

species such as maize and sorghum (~70%; as part of the Panicoideae) or between Triticeae species and oat (~60%; as part of the Poideae). However, these data suggested extensive rearrangements between the grass genomes and brought a note of caution into the concept of using small grass genomes (rice or sorghum) as a proxy for more complex genomes (maize or wheat) (Gaut 2002).

17.2.2 Sequence Based Macrocolinearity Studies

In the past 6 years, the release of the rice genome sequences (Feng et al. 2002; Goff et al. 2002; Sasaki et al. 2002; Yu et al. 2002; The Rice Chromosome 10 Sequencing Consortium 2003; International Rice Genome Sequencing Project 2005; The Rice Chromosome 3 Sequencing Consortium 2005) and the development of large EST collections from other grass species including the Triticeae, have allowed more in depth analyses of the colinearity between the Triticeae and the other grass genomes. For wheat and barley, the International Triticeae EST Cooperative (ITEC; <http://wheat.pw.usda.gov/genome/>) has permitted to increase the number of ESTs present in the public databases from 6 in 1998 to more than 1.5 million (1,084,449 for wheat; 478,734 for barley) to date (http://www.ncbi.nlm.nih.gov/dbEST/dbEST_summary.html). Large programs aiming at either globally assigning (in genetic or deletion bins) or precisely mapping Triticeae ESTs have resulted in the mapping of 7,107 EST singletons (16,099 loci) into a chromosome bin map using sets of euploid and aneuploid lines in wheat (Qi et al. 2004) and in the development of a genetic map with more than 1,000 gene loci in barley (Stein et al. 2007).

In silico comparative analyses were then performed by comparing the sequences of EST markers mapped in Triticeae species to each other and to the rice genome sequence to study macrocolinearity at high resolution. Because these comparisons are based on sequence alignments and because in most of the cases it is difficult to infer orthologous and paralogous relationships from sequence alignments, statistical analyses are required to objectively evaluate whether the association between two or more genes in the same order on two chromosomal segments occurs by chance or reveals significant colinearity. Several softwares such as LineUP (Hampson et al. 2003), ADHoRE (Automatic Detection of Homologous Regions, Vandepoele et al. 2002) and FISH (Fast Identification of Segmental Homology, Calabrese et al. 2003) have been developed for this purpose. A number of programs such as Cmap (Fang et al. 2003) and websites such as Gramene (<http://www.gramene.org/cmap/>; Jaiswal et al. 2006) have also been developed to visualize the sequence-based colinearity between the grass genomes obtained from these comparisons. Outputs of the sequenced-based macrocolinearity between the rice genome sequence and 9,332 wheat genetic marker sequences can be viewed at <http://www.tigr.org/tdb/synteny/wheat/description.shtml>. While these websites provide “user friendly” graphical displays of macrocolinearity, they rely on data obtained with low

stringency alignment criteria and without statistical validation. In addition, they do not take into account the density and location of conserved genes to identify precisely paralogous and orthologous regions and therefore they, generally, overestimate colinearity between different segments of the genomes.

In wheat, Sorrells and collaborators undertook the first comparative genome wide studies using 4,485 wheat ESTs mapped in deletion bins against the rice genome (Sorrells et al. 2003; Sorrells 2004). This increased the resolution of comparative mapping with rice by 25–30 fold compared to the previous RFLP-based assessments and allowed to specify the degree of conservation between each of the orthologous chromosomes. The results showed that wheat chromosome group 3 is the most conserved while wheat chromosome group 5 is the least conserved compared to rice (La Rota and Sorrells 2004). Studies focusing on single chromosome groups or regions were also performed for rice chromosome 3 compared to the wheat and maize genomes. They indicated high colinearity, further supporting the use of rice as a model for comparative studies among the cereals (Buell et al. 2005; The Rice Chromosome 3 Sequencing Consortium 2005). In contrast, in a comparative study between rice chromosome 11 and wheat, Singh et al. (2004) concluded for a lack of extensive colinearity and suggested that many insertions, deletions, duplications and translocations complicate synteny evaluation in cereals. In an extensive colinearity study using RFLPs, candidate genes and EST sequences from the short arm of wheat chromosome 1A and rice chromosome 5S, Guyot et al. (2004) found frequent disruptions of the markers order resulting in a mosaic conservation of genes in this region. In barley, Stein et al. (2007) produced an integrated transcript map based on 1,032 EST-based markers with one marker every 0.9 cM in average. Forty-six percent (475) of the markers were assigned to orthologous linkage groups in rice. Chromosome pairs such as 3H (barley)/r1(rice) or 6H/r2 were collinear over their entire length while some barley chromosomes resemble a mosaic of individual rice chromosomes such as 5H/r3-r9-r12.

More recently, colinearity relationships between the Ehrhartoideae (rice), Triticeae (wheat, barley, rye) and the Panicoideae (sorghum, maize) chromosomes was reassessed using the latest genome and EST releases in these species as well as improved sequence alignment criteria and systematic statistical analysis (Salse et al. 2008) (Table 17.1).

Sequence comparisons between the 42,654 genes annotated in rice (ftp://ftp.tigr.org/pub/data/Eukaryotic_Projects/o_sativa/annotation_dbs/pseudo-molecules/version_4.0/) and the 6,426 mapped wheat ESTs revealed 13 blocks of colinearity that represent 83.1% and 90.4% of the rice and wheat genomes, respectively. They correspond to the following chromosome pairs: w1-r5, w1-r10, w2-r4, w2-r7, w3-r1, w4-r3, w4-r11, w5-r3, w5-r9, w5-r12, w6-r2, w7-r6 and w7-r8 (Table 17.1) (Salse et al. 2008). A comparison of the linear order of the rice genes with the position of the orthologous ESTs in the wheat deletion bins indicated that for 27.2% of them, the identified wheat ortholog is not located in the orthologous wheat deletion bin thereby indicating additional

Table 17.1 Colinearity between the Ehrhartoideae (rice), Triticeae (wheat, barley, rye) and Panicoideae (sorghum, maize) chromosomes (updated from Salse et al. 2008)

<i>Ehrhartoideae</i> <i>Rice</i> <i>chromosomes</i>	<i>Triticeae</i>		<i>Panicoideae</i> <i>Sorghum</i> <i>chromosomes</i>	<i>Maize</i> <i>chromosome</i>
	<i>Wheat (W)</i>	<i>Barley (H)</i>		
	<i>Rye (R)</i> <i>chromosomes</i>			
r5	1		sG	m6
				m8
r1	3		sA	m8
				m3
r3	4		sC	m1
	(5)			m5
				m9
r7	2		sB	m7
				m2
r10	1		sC	m1
				m5
				m9
r11	4		sH	m4
				m2
r12	5		sE	m10
				m3
				m1
				m4
r8	7		sJ	m4
				m1
				m6
				m10
r9	5		sB	m7
				m2
r2	6		sF	m4
				m5
r4	2		sD	m2
				m10
r6	7		sI	m9
				m6
				m5

rearrangements within orthologous regions between rice and wheat (Salse et al. 2008). These results showed that even if the current public set of mapped wheat EST presents some limits in comparative analysis, since the linear order of the ESTs within a wheat deletion bin is not known, rearrangements can be identified and the evaluation of colinearity between wheat and rice can be improved through an accurate assessment of the size and position of the orthologous regions.

Thus, genome wide sequence based comparisons have revealed additional chromosomal rearrangements between the grass genomes than those previously

reported based on RFLP analyses. It provided a more complex picture of the orthologous relationships between these genomes and led to a first revision of the “concentric crop circles” in 2005 (Devos 2005). Future large EST mapping or genome sequencing projects in grasses including brachypodium, wheat and barley (see Chapters 16 and 24) will help to continue refining the degree of colinearity between the grass genomes as well as perform high resolution colinearity studies within the Triticeae tribe.

17.3 Comparative Genomics at the “Locus-Based” Level: Microcolinearity

Increasing evidence that rearrangements limit the extend of colinearity between the Triticeae genomes and model genomes led Triticeae geneticists to develop genomic resources e.g. large insert BAC libraries from their own species (wheat, barley and rye) to perform map-based cloning and study genome structure and evolution. Technological improvements have allowed the construction of BAC libraries with a sufficient number of clones to provide a reasonable coverage even from large and complex genomes such as those of wheat and barley (Chalhoub et al. 2004). In addition, for polyploid species such as wheat, advances in flow sorting techniques (Kubalakova et al. 2002) have allowed the isolation of DNA in sufficient amounts and quality to construct BAC libraries from single chromosomes or chromosome arms (Safar et al. 2004; Janda et al. 2004, 2006) and the development of chromosome genomics (see Chapter 10). A number of BAC libraries are now available for wheat, barley and rye (see Table 9.1 in Chapter 9) to perform microcolinearity studies at different levels.

17.3.1 Interspecific Comparative Studies: Looking at 50–70 MY of Speciation

One of the first microcolinearity study involving Triticeae species was performed at the Shrunk2/Anthocyaninless1 (*sh2/a1*) orthologous locus that was originally studies in maize, sorghum and rice (Chen et al. 1997, 1998). Despite large differences in the length of the intergenic regions in maize compared to rice and sorghum and a tandem duplication of one gene (*A1*) in sorghum, the linear order of the four genes (*Sh2*, *X1*, *X2* and *A1*) present at this locus was remarkably conserved between the three species. In contrast, in the Triticeae, colinearity was limited to the conservation of the *Sh2* and *X1* genes on chromosome 1L whereas the two other genes, *X2* and *A1*, were found on a non orthologous chromosome (3L). This indicated that numerous rearrangements including genes translocation have occurred at the locus since the divergence between the Triticeae and the other grasses (Li and Gill 2002). Since

these first studies, several other microcolinearity studies involving Triticeae species have been performed at different loci carrying genes involved in disease resistance (e.g. *Lrk*, *Lr34*, *Pm3*, *vrs1* *Rph7*), plant development (e.g. *Vrn1-3*, *PhdH1*, *Ph1*), and quality (e.g. *Ha*, *Glu*) (Table 17.2).

Table 17.2 Inter and intra specific microcolinearity studies involving triticeae genomes. List of the various loci that have been compared at the sequence level through BAC sequencing between different cereal species including at least one Triticeae genome (bold). The asterisk indicates intraspecific comparisons

Locus	Compared plant species	References
<i>Lrk</i>	Wheat, barley, maize, rice	<i>Feuillet and Keller 1999</i>
<i>Rph7</i>	Barley, rice	<i>Brunner et al. 2003</i>
<i>Vrn1</i>	Wheat, barley, sorghum, rice	<i>Dubcovsky et al. 2001</i> <i>Ramakrishna et al. 2002</i>
<i>Vrn2</i>	Wheat, barley, rice	<i>Yan et al. 2004</i>
<i>Vrn3</i>	Wheat, barley, rice	<i>Yan et al. 2006</i>
<i>sh2/a1</i>	Wheat, maize, sorghum, rice	<i>Chen et al. 1997</i> <i>Chen et al. 1998</i> <i>Li and Gill 2002</i> <i>Bennetzen and Ma 2003</i>
<i>Ha</i>	Wheat*, barley, rice	<i>Caldwell et al. 2004</i> <i>Chantret et al. 2005</i>
<i>Phd-H1</i>	Barley, rice	<i>Dunford et al. 2002</i>
<i>Glutenin</i>	Wheat*, rice	<i>Wicker et al., 2003</i> <i>Gu et al. 2004</i>
<i>Pm3</i>	Wheat*, rice	<i>Wicker et al. 2007</i>
<i>Lr34</i>	Wheat, rice, Brachypodium	<i>Bossolini et al. 2007</i>
<i>BCD135, RZ567</i>	Barley, rice	<i>Park et al. 2004</i>
<i>Ph1</i>	Wheat, rice, Brachypodium	<i>Griffith et al. 2006</i>
<i>vrs1</i>	Barley, rice	<i>Pourkheirandish et al. 2007</i>

They all confirmed that many small-scale genic rearrangements, such as single or multiple gene insertions and/or deletions, tandem duplications, inversions, and translocations that were previously overlooked by comparative genetic mapping have occurred during the evolution of the cereal genomes including the Triticeae (for reviews see Feuillet and Keller 2002; Bennetzen and Ramakrishna 2002; Devos 2005; Salse and Feuillet 2007). Depending on the chromosomal location and type of locus, the extent of conservation can vary from single gene differences to complete disruption of colinearity due to translocations. Disease resistance (R) loci are particularly prone to rearrangements compared to loci involved in the primary metabolism. The disruption of colinearity between the grass genomes at R loci was already observed by comparative mapping in rice, barley and foxtail millet using resistance gene analogs as RFLP probes (Leister et al. 1998). Very recently, Wicker et al. (2007) performed a comparative and phylogenetic analysis between sequences of the wheat powdery mildew disease resistance *Pm3* genes originating from the A genome at different polyploidy levels and their orthologs in rice. The results showed that

all the wheat *Pm3*-like genes cluster together within a subgroup of the rice homologs indicating that the *Pm3* gene family has evolved multiple subfamilies well before the divergence of rice and wheat (Wicker et al. 2007). Estimates of divergence times and transposable-element insertions showed that the *Pm3* locus in wheat has undergone more drastic changes in its recent evolution than its ortholog in rice thereby indicating that loci containing homologous resistance gene analogs can evolve at highly variable speeds in different species.

These comparisons have been very helpful to identify some of the mechanisms involved in the rearrangements that have shaped the grass genomes during their evolution. It is clear now that transposable elements (TEs) and in particular retrotransposons have played a major role in the expansion of the large genomes of the Triticeae through nested insertions and that numerous small deletions caused by unequal homologous recombination and illegitimate recombination have counteracted this expansion (Bennetzen et al. 2005). Even though Triticeae species or more generally grass genomes do not share common TEs at orthologous positions, several studies have demonstrated that repeated element families are shared between cereal genomes. For example, Lisch et al. (2001) and Lisch (2002) demonstrated that the Mutator-like (MuLE) transposons initially identified in maize are ubiquitous and diverse in grasses. Moreover, they identified cDNAs for this element in rice and wheat, suggesting the presence of active MuLEs in these species. The CACTA transposons family has also been investigated in the Triticeae. The identification by Wicker et al. (2003) of CACTA elements in wheat as well as the identification of similar elements in sorghum and rice suggests that CACTA transposons contributed significantly to grass genomes organisation and evolution. Finally, a phylogenetic analysis of the retrotransposon family, which composes most of the Triticeae genomes, has been performed by Wicker et al. (2007) using 599 copia-retrotransposons identified in barley, wheat, rice and Arabidopsis. This led to the detection of 6 ancient lineages that existed before the divergence of monocots and dicots.

TEs have clear ancient insertion patterns in cereals and more precisely in the Triticeae genomes. For example in barley, BARE (BARley REtrotransposons) elements represent up to 10% of the genome (Suoniemi et al. 1996a,b; Waugh et al. 1997; Chapter 14). However, due to their high transposition activity and rapid turnover, TE in orthologous position can only be detected in the case of recent insertions and in the context of intraspecific comparisons.

17.3.2 Intraspecific Comparisons: Microcolinearity Studies Within Few MY of Speciation

With the development of BAC libraries from different subspecies in rice and from species at different ploidy levels in wheat, intraspecific microcolinearity studies have been performed within divergence time smaller than 5 million years. Several studies have compared sequences of the homoeologous A, B,

and D genomes of wheat that are estimated to have diverged from a common ancestor between 2.5 and 4.5 million years ago (Huang et al. 2002). BAC sequences originating from two different haplotypes identified at the disease resistance locus *Lr10* have been compared in diploid (*T. monococcum*, A^m genome), tetraploid (*T. durum*, AB genome) and hexaploid wheat (*T. aestivum*, ABD genome). Insertions as well as deletions and unequal crossing over between transposable elements have reduced the overall percentage of sequence conservation between the three orthologous regions to 33% and very few elements were conserved in the intergenic regions even within the same haplotype. A good conservation of the gene content and order was found between the diploid and tetraploid sequences that belong to the same haplotype but a large rearrangement involving a deletion followed by a large inversion was observed in the second haplotype in hexaploid wheat. This work also allowed to estimate the divergence time between the A genomes of wheat to 2 MY demonstrating the interest of performing intraspecific comparative studies (Isidore et al. 2005).

Comparative sequencing was also performed at *Ha* locus that controls grain hardness in wheat. Orthologous BACs were compared in *Triticum aestivum*, *Triticum durum*, the diploid relatives *Triticum monococcum* and *Aegilops tauschii*, and rice (Chantret et al. 2005). Rearrangements, such as transposable element insertions, sequence deletions, duplications, and inversion involving illegitimate recombination were shown to be responsible for the major differences observed between the same genomes at different ploidy levels (Fig. 17.2). The results showed that *Pin* genes (*Pina*, *Pinb*) which are absent in rice, are conserved in wheat diploid progenitors e.g. on the D genome of *Ae. tauschii* and the A^m genome of *T. monococcum* but lack in the A and B genomes of tetraploid wheat *Triticum turgidum* (Fig. 17.2).

These sequence comparisons provided insight into the origin of the previously reported loss of *Pina* and *Pinb* genes which is related to grain hardness trait in tetraploid wheat. The *Pina* and *Pinb* genes were removed through large deletions resulting from illegitimate recombination mechanisms that have occurred independently in the A and B genomes following polyploidisation (Fig. 17.2). Finally, three studies have compared orthologous glutenin gene loci. The first two were performed between the A and B genomes of *T. durum*, and the D genome of *Ae. tauschii* (Gu et al. 2004) and between the homoeologous A genomes of *T. durum* and *T. monococcum* (Wicker et al. 2003). They indicated that microcolinearity is maintained between homoeologous wheat genomes, but that intergenic regions are not conserved due to rapid amplification/deletion of retroelements and illegitimate recombination. A more recent analysis that includes the three (A, B, D) sub-genomes of hexaploid wheat indicated a higher level of conservation in the intergenic regions between homoeologous genomes (Gu et al. 2006). Very recently, Wicker et al. (2007) performed a comparative analysis of BAC sequences originating from the A genomes of *T. monococcum*, *T. durum* and *T. aestivum* at the powdery mildew disease resistance locus *Pm3* in wheat. The results demonstrated extreme reshuffling and an almost complete disruption of orthology at the *Pm3* locus in wheat

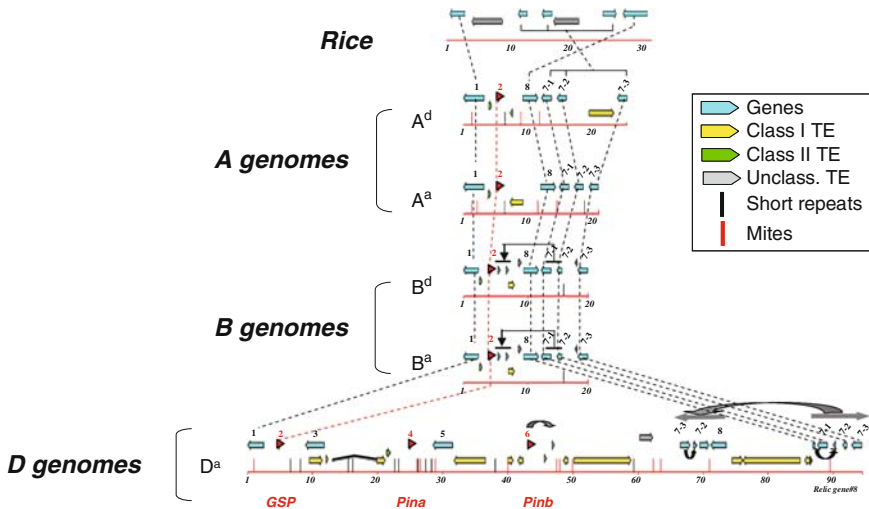


Fig. 17.2 Microcolinearity studies at the Hardiness locus in wheat (adapted from Chantret et al. 2005). Schematic representation of BAC sequence comparisons at the wheat Ha locus from the A (A^a : *T. aestivum*; A^d : *T. durum*), B (B^a : *T. aestivum*; B^d : *T. durum*) and D (D^a : *T. aestivum*) wheat and rice genomes. Genes (CDS) (light blue), class I TEs (blue), class II TEs (dark blue), unclassified elements (gray), MITEs (red arrows), and repeats (black arrows) are indicated. Orthologous CDS between the different genomes are linked by dashed bars whereas CDS duplications and deletion events are indicated by arrows. The GSP, Pina and Pinb genes are numbered respectively as gene 2, 4, 6 (See Color Insert)

within a short evolutionary time. Basically, the conservation between the homoeologous loci was restricted to small regions containing the sequence used to screen the BAC library and a few other short segments. Crossing over between repetitive elements led to the deletion of genic regions and their replacement by repetitive elements in *T. aestivum* compared to *T. monococcum* (Wicker et al. 2007). Thus, comparative analysis between homoeologous genomes led to the identification of major mechanisms involved in both expansion and reduction of the wheat genomes. They suggest that TEs have been very active since the divergence of the A, B, and D genomes as well as after polyploidization and that illegitimate DNA recombination, leading to various genomic rearrangements, is one of the major evolutionary mechanisms in these genomes as observed in general in plant genomes.

17.3.3 Intravarietal Comparisons: Microcolinearity Studies Within Few 10,000 Years of Speciation

The observed absence of colinearity at the intraspecific level between recently diverged species raised the question of sequence rearrangements within

different lines or varieties. The first evidence for violation of microcolinearity between different lines within a species was found in maize when BAC sequences from the bronze (*bz1*) locus were compared in two maize inbred lines (McC and B73). The results revealed as dramatic differences between the two lines as those observed between orthologous loci in two different species (Fu and Dooner 2002; Lai et al. 2005). The absence of microcolinearity did not concern only the length and composition of the intergenic regions, but also affected the gene density and content. Four out of ten genes present in McC were absent in B73. These findings were then further confirmed by the comparative analyses of other loci in different inbreds e.g. the zein storage protein gene cluster *z1C-1* in B73 and BSSS53 (Song and Messing 2003) and four allelic chromosomal regions in Mo17 and B73 (Brunner et al. 2005). This latter study showed that almost 50% of the total sequence analyzed was not shared between the two inbreds. Most of it consisted of LTR-retrotransposons and other mobile elements but there were also considerable differences in the genic sequences. In total, 23 out of 68 putative genes (34%) were present only in either Mo17 or B73. In contrast to the *z1C-1* locus where half of the non-shared sequence originated from extensive local duplications, the non-shared sequences corresponded to clusters of genes fragments. Interestingly, in contrast to the shared genes, the non-shared genes were not present at colinear positions in rice suggesting that they likely originate from insertions rather than deletions. Latter on, Morgante et al. (2005) identified the non shared pseudo-gene clusters as part of non-autonomous Helitrons, a new type of eukaryotic transposable elements. These transposons appear to have copied and incorporated genic segments from different genomic locations of the host, clustered them together and duplicated these arrangements via a copy-past transposition mechanism to non-allelic loci across the maize genome (for references see Morgante et al. 2005).

Few such studies have been performed in the Triticeae species so far. In barley, more than 300 kb of sequence spanning the *Rph7* leaf rust disease resistance gene has been compared between two cultivars (Scherrer et al. 2005). Colinearity was restricted to five genic and two intergenic regions representing less than 35% of the two sequences. In each interval separating the conserved regions that were mainly found around genes, the number and type of repetitive elements were completely different. A single gene that was identified later as a helitron (C. Feuillet, personal communication) was absent in one cultivar. In both cultivars, the non-conserved regions consisted of ~53% repetitive sequences mainly represented by long-terminal repeat retrotransposons that have inserted less than 1 million years ago. PCR-based analysis of intergenic regions at the *Rph7* locus and at three other independent loci in 41 *H. vulgare* lines indicated rapid and recent divergence at homologous loci in the cultivated barley genome (Scherrer et al. 2005). The rearrangements observed in barley were less dramatic than those found between maize inbreds as well as those observed between rice subspecies suggesting that maize has a highly unstable genome compared to the other grasses (Fu and Dooner 2002).

No comparative analysis has been performed yet between different wheat varieties of the same ploidy level. However, BAC libraries from different hexaploid wheat cultivars are available now (cf. Chapter 9) and comparative studies are underway (http://www.intl-pag.org/14/abstracts/PAG14_W30.html). It will be interesting to compare the rate and mechanisms of evolution at a similar time scale in barley and wheat, two species that are closely related but have very different population histories.

17.4 Duplications in the Triticeae Genomes

In addition to the assessment of colinearity between the genomes, comparative analyses can reveal ancestral genome duplications. Early studies with the first generation of molecular markers indicated the presence of duplicated loci on the genetic maps in different cereals, including the Triticeae, suggesting ancestral genome duplications and polyploidization events in the history of grasses. RFLP and isozymes studies in the early 90s had already suggested that maize chromosomes share duplicated segments (Wendel et al. 1989; Ahn and Tanksley 1993; Gaut 2001). Whole duplication of the maize genome through allotetraploidization was identified and characterized further through the evolutionary analysis of duplicated genes (Gaut and Doebley 1997) and by interspecific comparisons between orthologous loci in rice, sorghum, and maize (Swigonová et al. 2004). In rice, early RFLP mapping studies suggested that chromosomes 1 and 5 (Kishimoto et al. 1994) as well as chromosomes 11 and 12 (Nagamura et al. 1995) contain ancient duplicated regions. The release of the genome sequence drafts from japonica and indica rice subspecies allowed whole genome sequence comparisons and further characterization of duplications in rice (Vandepoele et al. 2003; Paterson et al. 2004; Guyot et al. 2004; Yu et al. 2002, 2005; IRGSP 2005; Wang et al. 2005). These studies revealed a whole genome duplication (WGD) that occurred between 53 and 94 MYA (i.e. before the divergence of the cereal genomes), a recent segmental duplication between chromosomes 11 and 12 and numerous individual gene duplications. Duplication in the Triticeae genomes has been addressed very recently through the analysis of mapped ESTs in wheat and barley and their comparisons with the rice genome. In wheat, we recently performed intraspecific sequence comparisons using 6,426 mapped wheat ESTs and identified 10 duplication blocks covering 67.5% of the genome (Salse et al. 2008). Further comparisons with 29 identified duplicated regions in rice revealed that seven of the intra-specific duplications are conserved at orthologous positions between rice and wheat. These ancestral shared duplications were found on the following chromosome pair combinations: w1–w2/r5–r4, w1–w3/r5–r1, w1–w4/r10–r3, w2–w4/r7–r3, w2–w7/r4–r8, w5–w7/r9–r8 and w6–w7/r2–r6 (Fig. 17.3). Altogether, they represent 68.3% of the rice genome and 65.9% of the wheat genome.

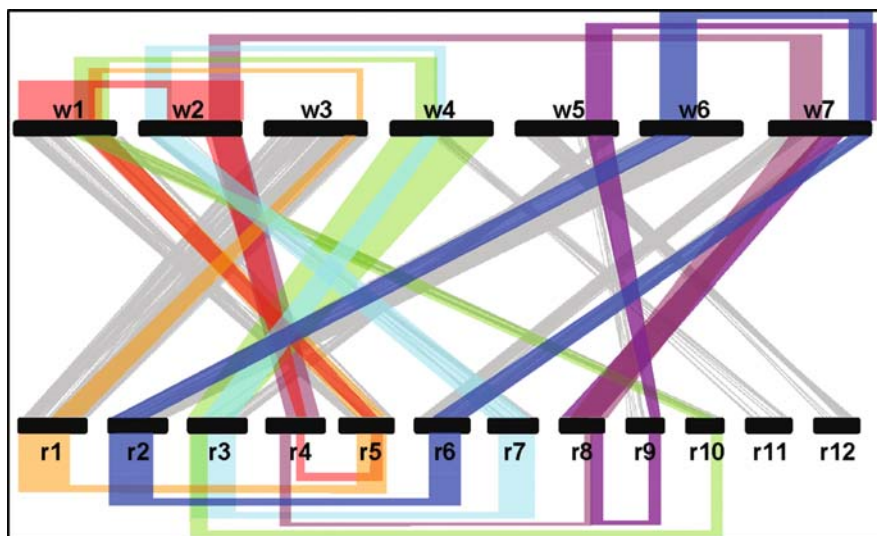


Fig. 17.3 Seven duplications shared between rice and wheat (updated from Salse et al. 2008). The paralogous regions are represented with the same colors in wheat (*top*) and rice (*bottom*) as well as the corresponding orthologous regions identified between the two sets of chromosomes (*center*). The color code is: w1–w2/r5–r4 (red); w1–w3/r5–r1 (orange); w1–w4/r10–r3 (green); w2–w4/r7–r3 (light blue); w2–w7/r4–r8 (pink); w5–w7/r9–r8 (violet); w6–w7/r2–r6 (dark blue); rice–wheat orthologs not involved in shared duplications are indicated in grey (See Color Insert)

In barley, Stein et al. (2007) assigned 475 EST markers to syntenic linkage groups of rice and performed dot plot comparisons between the barley and rice chromosomes. Using these data, the authors analysed duplications on chromosomes 2H and 6H in more details through comparisons with the rice genome duplications. The results showed that the rice duplication r4–r2 is conserved in barley between chromosomes 2H–6H (Stein et al. 2007) supporting the idea that ancestral duplications are shared between the Triticeae and rice genomes.

To further analyse the origin and evolution of the shared duplications, we recently combined the data obtained in wheat and rice with previous comparative analyses performed between rice and maize (Salse et al. 2004; Wei et al. 2007) and between rice and sorghum (Paterson et al. 2004). Analysis of the conservation pattern of shared duplication in the four species led to propose a model for the grass genome evolution from an ancestor with 5 chromosomes that underwent a whole genome duplication (WGD), 50–70 MYA followed by two interchromosomal translocations and fusions that resulted in a $n=12$ intermediate ancestor (Fig. 17.4) (Salse et al. 2008). In this model, rice would have retained the 12 original chromosome number whereas it has been reduced in the other cereal genomes. In wheat, 5 chromosomal fusions resulted in an ancestral wheat genome with $n=7$ chromosomes (Fig. 17.4). Thus, the 10 duplicated regions observed in the wheat genome reflect the ancestral WGD

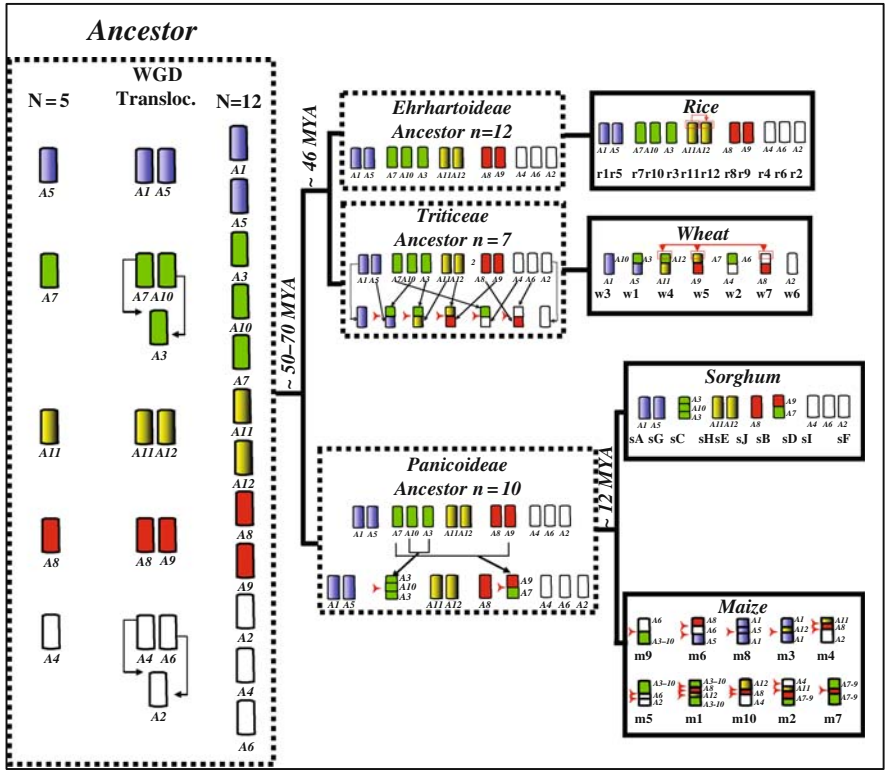


Fig. 17.4 Model for the structural evolution of the rice, wheat, sorghum, and maize genomes from a common ancestor with $n=5$ chromosomes (adapted from Salse et al. 2008). Chromosomes are represented with color codes to illuminate the evolution of segments from a common ancestor with 5 chromosomes. The 5 chromosomes are named accordingly to the rice nomenclature. Where possible, divergence times are indicated based on previous estimates by Gaut (2002), Swigonová et al. (2004) and Paterson et al. (2004) (See Color Insert)

and 3 additional segmental duplications (SD) that have occurred since the chromosomal fusions.

The results obtained in barley (Stein 2007) and the high colinearity observed between the Triticeae genomes suggest that this pattern reflects the evolution of the ancestral Triticeae genome (Fig. 17.4). For maize and sorghum, our findings and model is in complete agreement with the recent analysis of Wei et al. (2007) who showed that both have evolved from an ancestral genome with 12 chromosomes after 2 chromosomal fusions resulting in an intermediate Panicoideae ancestor with $n=10$ chromosomes (Fig. 17.4). Then, the two species evolved independently from this ancestor. While the sorghum genome structure remained similar to the ancestral genome, maize underwent an allotetraploidization (Gaut and Doebley 1997; Swigonová et al. 2004) that resulted in an intermediate with $n=20$ chromosomes (Salse et al. 2008). Rapid and extensive

chromosome reshuffling led then to a mosaic maize genome with 10 chromosomes (Fig. 17.4).

Even if comparative genetic mapping and EST sequence comparisons can be used to identify ancient duplications, there is nothing more powerful than a genome sequence to study the origin and mechanisms of ancient genome duplications. Thus, there is no doubt that the projects that are underway for sequencing the wheat and barley genomes (cf. Chapter 24) will help to extend these studies and shed new light into the duplication events that have affected the Triticeae genomes evolution as well as deliver very efficient tools and data to support the isolation of genes of agronomic interests in these species.

17.5 Comparative Genomics as Tool for Gene Discovery and Marker Development

Comparative genomic studies have led to a better knowledge of the level of conservation between the cereal genomes and to the generation of genomic tools that can now be used to define efficient strategies for genetic studies and gene isolation in these genomes and in particular those of the Triticeae for which no physical map and no genome sequences are available yet.

17.5.1 Colinearity-Based Gene Cloning in Triticeae

In some cases, the conservation of sequence at orthologous positions between the genomes can reflect the conservation of a gene with a similar function between species. Early comparative genetic studies using RFLP had identified a number of genes and quantitative traits loci (QTL) for developmental and domestication traits, e.g. shattering, plant height, vernalisation, flowering time, row number and kernels per row, at orthologous positions in cereal genomes (Lin et al. 1995; Paterson et al. 1995; Bailey et al. 1999). The concomitant discovery of colinearity between rice and the other cereals opened up very exciting perspectives in using the rice genome to support positional cloning of genes from the other genomes in a so called “cross genome map-based cloning” approach, even before the rice sequence was completed (Kilian et al. 1997). The best example of colinearity in gene type and function and in the efficient use of rice for direct gene cloning in Triticeae was illustrated by the isolation of *Rht-1* gene in wheat (Peng et al. 1999). In the last years, the isolation of genes by map-based cloning in barley and wheat has also revealed examples of conservation between genes at orthologous positions in cereals (Table 17.2). For example, the wheat vernalisation gene, *Vrn1* (Yan et al. 2003) and the barley photoperiod PPD-H1 gene both have orthologous genes in rice (Turner et al. 2005). In regions where microcolinearity is high, candidate genes can be directly identified from the rice sequence. This has been successfully used to support the

isolation of the powdery mildew resistance gene *Ror2* (Collins et al. 2003) and the *sw3* dwarfism gene in barley (Gottwald et al. 2004). In other cases, similar functions do not seem to be associated with similar genes. For example, in a study comparing QTL for heading time in rice and barley, Griffiths et al. (2003) showed that in rice a number of QTL belong to the CONSTANS gene family but that in barley none of the homologous CONSTANS genes are associated with any of the known QTL for flowering time. Thus, generally genes and QTL involved in developmental processes and that have been selected during domestication show a good conservation between cereal genomes and rice genes are good candidates for direct gene isolation.

In contrast, other types of genes do not show colinearity between the grass genomes. Indeed, there is no example of colinearity retained for disease resistance (R) genes between grass genomes and so far, map-based cloning of R genes in Triticeae was not significantly profiting from the rice genome information. The non-syntenic location of R genes in the cereals was already described 10 years ago through a comparative genetic analysis of resistance gene analogs in barley, rice and foxtail millet (Leister et al. 1998). In many cases, the attempts of using colinearity with rice for isolating R genes have revealed the limits of collinearity between rice and the other cereal genomes including the Triticeae. The first note of caution was provided with the barley stem rust resistance gene *Rpg1* map based cloning project. Despite a certain degree of colinearity retained at the orthologous locus in rice (Kilian et al. 1997), no orthologous gene was present in the rice genome and map-based cloning of *Rpg1* had to be achieved in barley (Brueggeman et al. 2002). In some cases, such as the leaf rust *Lr10* and the powdery mildew *Pm3* fungal disease R genes on wheat chromosome 1AS, the rice genome contains genes homologous to the wheat candidate genes but at non-orthologous positions, indicating massive genome rearrangements (Guyot et al. 2004). Both R genes were also cloned using alternative strategies (Stein et al. 2000; Feuillet et al. 2003; Yahiaoui et al. 2004; Chapter 12). The only known exception to this lack of colinearity between disease R genes has been reported by Chen et al. (2005) who showed that a QTL conferring resistance to the blast fungus *Magnaporthe grisea* is conserved at orthologous positions and with the same race specificity in rice and barley.

Even if the gene is not present at its orthologous position in rice, flanking genes are often conserved enough to provide a source of markers than can be used to saturate the target region in the Triticeae genomes. For example, rice ESTs were used to reduce the genetic interval around the barley disease R loci *Rpg1* and *Rph7* to a density that allowed initiation of chromosome walking in barley (Brueggeman et al. 2002; Brunner et al. 2003). There are now many additional examples of the use of rice EST derived markers to saturate genetic regions in Triticeae and this approach is now routinely used in laboratories that are involved in Triticeae gene cloning worldwide (Yan et al. 2003, 2004, 2006; Sutton et al. 2003, 2007; Griffiths et al. 2006). In many cases, barley EST markers were also very helpful to saturate wheat loci. Recently, a new model species, *Brachypodium*, has been proposed (Draper et al. 2001; Vogel et al.

2006; Hasterok et al. 2006) for temperate cereals such as wheat and barley (see Chapter 16). It was successfully used in combination with rice to isolate *Ph1*, one of the key gene controlling chromosome pairing in polyploid wheat (Griffiths et al. 2006; Chapter 8) and increase marker density at the *Lr34* locus in wheat (Bossolini et al. 2007).

17.5.2 Comparative Genomics Supports Gene Annotation and Marker Development

Complete genome sequences provide the basis for understanding the gene structure and function within species. As genes are the most conserved features between genomes, the availability of a genome sequence can greatly help to predict genes from other genomes. Even between distantly related genomes such as the one of rice and *A. thaliana*, which ancestors diverged 200 million years ago and do not show extensive macrocolinearity, a large number of genes have been conserved (Salse et al. 2002). Thus, the rice genome sequence represents a unique tool to support gene annotation in other cereal genomes, a critical issue in view of the future sequencing of the Triticeae genomes (Chapter 24). While waiting for the production and release of large public FLcDNAs sets in wheat and barley, the availability of more than 28,000 rice full length cDNAs (The Rice Full-Length cDNA Consortium 2003) is particularly useful for gene annotation of Triticeae genomic sequences and can help to validate intron/exon boundaries and train gene predictors. The brachypodium genome sequence that is planned for the end of 2008 (<http://www.brachypodium.org/> and Chapter 16) will also be of great support and brachypodium sequences have already proven to be useful to improve rice gene annotation (Bossolini et al. 2007). Conversely, the alignment of EST from Triticeae species with the rice and brachypodium genome sequences can also help to predict new genes from these species.

The accurate identification of intron/exon boundaries is also very interesting for the development of new markers. Indeed, SNP frequencies are higher in introns than in exons and the possibility to design PCR primers amplifying intronic sequences can greatly improve polymorphism detection in species, such as wheat, that suffer from a chronic lack of polymorphism. This concept has been applied in pearl millet by Bertin et al. (2005) who have aligned millet ESTs against the rice gene sequences to predict the location of introns and amplify products across them followed by detection by Single strand Conformational Polymorphism (SSCP). The SSCP-SNP marker technique has a great potential for the development of COS (Conserved Orthologous Set) markers for comparative mapping in cereals if comparisons are performed between sequences from many different species and can be used to define perfect match primers. We have recently defined 695 COS markers for wheat mapping and QTL detection based on comparative analyses between rice and wheat. They show up to 50% of polymorphism between elite lines and are currently used for fine mapping in map based cloning projects (J. Salse, personal communication).

17.6 Summary and Outlook

The past decade has seen a revolution in Triticeae structural and functional genomics and has demonstrated the power of comparative studies in these economically essential crop species. Comparative studies led to improved genetic and physical maps, the development of large sets of accurate markers for breeding and the map-based isolation of the first genes of agronomic interest in wheat and barley. It also provided insight into the evolution of the grass genomes, unravelling some of the major mechanisms that have shaped their evolution within a 50–70 million years of speciation. Until recently, with only the rice genome sequence available, the power of comparative genomics in grasses has been limited. However, the situation is about to change dramatically with additional grass genomes sequences underway: maize (<http://www.genome.arizona.edu/>), sorghum (<http://www.phytozome.net/>), brachypodium (<http://www.brachypodium.org/>) and foxtail millet (<http://www.jgi.doe.gov/>), and the perspective of the barley and wheat genome sequences in the next decade (see Chapter 24). With these new resources, comparative genomic studies between cereal genome sequences will deliver additional information about plant genome evolution and better tools for crop improvement. As in animal comparative genomics, it will be possible to identify elements that have been conserved during the evolution and have functional significance. In mammals, comparisons of human, mouse and rat genomes showed that about 3% of the genome corresponding to non protein coding sequences are ultra conserved across genomes and have been under purifying selection (Bejerano et al. 2004). Conserved non coding regions (CNS) i.e. conserved sequences located in the non coding regions of genes (introns or upstream regulatory sequences) have been surveyed in cereals (maize vs. rice) and mammals (human vs. mouse) by Freeling and collaborators (Kaplinisky et al. 2002; Inada et al. 2003). They showed that CNSs are more abundant in regulatory genes such as transcription factors and that despite similar divergence times from their common ancestors, grass genes have dramatically fewer (5- to 20-fold) and smaller CNSs than mammalian genes. One possible explanation is that in contrast to vertebrate genomes, plant genomes have been subjected to several whole rounds of whole genome and/or segmental duplications, and polyploidization events that have profoundly affected their organisation, the subfunctionalisation of duplicated genes leading to a greater loss of CNS per gene (Lockton and Gaut 2005). Future comparative genome sequencing will help to confirm these features and provide clues to understand the relationship between CNS, regulation and phenotypes. In addition, CNSs also represent great target for PCR primer binding sites that can be used to design a new generation of COS markers for high density mapping in the Triticeae genomes.

Since a decade and with a great contribution of the Triticeae species, comparative genomics in cereals has pioneered the field of plant comparative genomics. There is no doubt that with the ongoing efforts, comparative studies

in grasses will continue to provide invaluable information for a better understanding of the adaptation of plants to their environment and open new areas for breeding strategies, plant protection and conservation of biodiversity.

References

- Ahn, S., and Tanksley, S.D. (1993) Comparative linkage maps of the rice and maize genomes. *Proc. Natl. Acad. Sci. USA* 90, 7980–7984.
- Bailey, P.C., McKibbin, R.S., Lenton, J.R., Holdsworth, M.J., Flintham, J.E. and Gale, M.D. (1999) Genetic map locations for orthologous Vp1 genes in wheat and rice. *Theor. Appl. Genet.* 98, 281–284.
- Bejerano, G., Pheasant, M., Makunin, I., Stephen, S., Kent, W.J., Mattick, J.S. and Haussler, D. (2004) Ultraconserved elements in the human genome. *Science* 304, 1321–1325.
- Bennett, M.D. and Smith, J.B. (1976) Nuclear DNA amounts in angiosperms. *Philos. Trans. R Soc. Lond. B Biol. Sci.* 274933, 227–274.
- Bennetzen, J.L. and Ramakrishna, W. (2002) Numerous small rearrangements of gene content, order and orientation differentiate grass genomes. *Plant Mol. Biol.* 48, 821–827.
- Bennetzen, J.L. and Ma, J. (2003) The genetic colinearity of rice and other cereals on the basis of genomic sequence analysis. *Curr. Opin. Plant Biol.* 6, 128–133.
- Bennetzen, J.L., Ma, J. and Devos, K.M. (2005) Mechanisms of recent genome size variation in flowering plants. *Ann. Bot.* 95, 127–132.
- Bertin, I., Zhu, J.H. and Gale, M.D. (2005) SSCP-SNP in pearl millet – a new marker system for comparative genetics. *Theor. Appl. Genet.* 110, 1467–1472.
- Bossolini, E., Wicker, T., Knobel, P.A. and Keller, B. (2007) Comparison of orthologous loci from small grass genomes *Brachypodium* and rice: implications for wheat genomics and grass genome annotation. *Plant J.* 494, 704–717.
- Brueggeman, R., Rostoks, N., Kudrna, D., Kilian, A., Han, F., Chen, J., Druka, A., Steffenson, B. and Kleinhofs, A. (2002) The barley stem rust-resistance gene *Rpg1* is a novel disease-resistance gene with homology to receptor kinases. *Proc. Natl. Acad. Sci. USA* 99, 9328–9333.
- Brunner, S., Keller, B. and Feuillet, C. (2003) A large rearrangement involving genes and low copy DNA interrupts the microcolinearity between rice and barley at the *Rph7* locus. *Genetics* 164, 673–683.
- Brunner, S., Fengler, K., Morgante, M., Tingey, S. and Rafalski, A. (2005) Evolution of DNA sequence nonhomologies among maize inbreds. *Plant Cell* 17, 343–360.
- Buell, C.R., Yuan, Q., Ouyang, S., Liu, J. et al. (2005) Sequence, annotation, and analysis of synteny between rice chromosome 3 and diverged grass species. *Genome Res.* 15, 1284–1291.
- Calabrese, P.P., Chakravarty, S. and Vision, T.J. (2003) Fast identification and statistical evaluation of segmental homologies in comparative maps. *Bioinformatics* 19(Suppl 1), i74–i80.
- Caldwell, K.S., Langridge, P. and Powell, W. (2004) Comparative sequence analysis of the region harboring the hardness locus in barley and its colinear region in rice. *Plant Physiol.* 136, 3177–3190.
- Chalhoub, B., Belcram, H. and Caboche, M. (2004) Efficient cloning of plant genomes into bacterial artificial chromosome (BAC) libraries with larger and more uniform insert size. *Plant Biotech. J.* 2, 181–188.
- Chantret, N., Salse, J., Sabot, F., Rahman, S. et al. (2005) Molecular basis of evolutionary events that shaped the hardness locus in diploid and polyploid wheat species (*Triticum* and *Aegilops*). *Plant Cell* 17, 1033–1045.

- Chen, H., Wang, S., Xing, Y., Xu, C., Hayes, P.M. and Zhang, Q. (2005) Comparative analyses of genomic locations and race specificities of loci for quantitative resistance to *Pyricularia grisea* in rice and barley. *Proc. Natl. Acad. Sci. USA* 100, 2544–2549.
- Chen, M., SanMiguel, P., deOliveira, A.C., Woo, S.S., Zhang, H., Wing, R.A. and Bennetzen, J.L. (1997) Microcolinearity in sh2-homologous regions of the maize, rice, and sorghum genomes. *Proc. Natl. Acad. Sci. USA* 94, 3431–3435.
- Chen, M., SanMiguel, P. and Bennetzen, J.L. (1998) Sequence organization and conservation in sh2/a1-homologous regions of sorghum and rice. *Genetics* 148, 435–443.
- Collins, N.C., Thordal-Christensen, H., Lipka, V. et al. (2003) SNARE-protein-mediated disease resistance at the plant cell wall. *Nature* 425, 973–977.
- Devos, K.M., Atkinson, M.D., Chinoy, C.N., Francis, H.A., Harcourt, R.L., Koebner, R.M.D., Liu, C.J., Masojc, P., Xie, D.X. and Gale, M.D. (1993a) Chromosome rearrangements in the rye genome relative to that of wheat. *Theor. Appl. Genet.* 85, 673–689.
- Devos, K.M., Millan, T. and Gale, M.D. (1993b) Comparative RFLP maps of the homologous group2 chromosomes of wheat, rye and barley. *Theor. Appl. Genet.* 85, 784–792.
- Devos, K.M. and Gale, M.D. (1997) Comparative genetics in the grasses. *Plant Mol. Biol.* 35, 3–15.
- Devos, K.M. and Gale, M.D. (2000) Genome relationships: the grass model in current research. *Plant Cell* 12, 637–646.
- Devos, K.M. (2005) Updating the ‘crop circle’. *Curr. Opin. Plant Biol.* 8, 155–162.
- Draper, J., Mur, L.J., Jenkins, G., Ghosh-Biswas, C., Bablak, P., Hasterok, R. and Routledge, A.P.M. (2001) Brachypodium distachyon. A new model system for functional genomics in grasses. *Plant Physiol.* 127, 1539–1555.
- Dubcovsky, J., Luo, M.C., Zhong, G.Y., Bransteitter, R., Desai, A., Kilian, A., Kleinhofs, A. and Dvorák, J. (1996) Genetic map of diploid wheat, *Triticum monococcum* L., and its comparison with maps of *Hordeum vulgare* L. *Genetics* 1432, 983–999.
- Dubcovsky, J., Ramakrishna, W., SanMiguel, P.J., Busso, C.S., Yan, L., Shiloff, B.A. and Bennetzen, J.L. (2001) Comparative sequence analysis of colinear barley and rice bacterial artificial chromosomes. *Plant Physiol.* 1253, 1342–1353.
- Dunford, R.P., Yano, M., Kurata, N., Sasaki, T., Huestis, G., Rocheford, T. and Laurie, D.A. (2002) Comparative mapping of the Barley Ppd-H1 photoperiod response gene region, which lies close to a junction between two rice linkage segments. *Genetics* 161, 825–834.
- Fang, Z., Polacco, M., Chen, S., Schroeder, S., Hancock, D., Sanchez, H. and Coe, E. (2003) cMap: the comparative genetic map viewer. *Bioinformatics* 19, 416–417.
- Feng, Q., Zhang, Y., Hao, P., Wang, S. et al. (2002) Sequence and analysis of rice chromosome 4. *Nature* 420, 316–320.
- Feuillet, C. and Keller, B. (1999) High gene density is conserved at syntenic loci of small and large grass genomes. *Proc. Natl. Acad. Sci. USA* 96, 8265–8270.
- Feuillet, C. and Keller, B. (2002) Comparative genomics in the grass family: molecular characterization of grass genome structure and evolution. *Ann. Bot.* 89, 3–10.
- Feuillet, C., Travella, S., Stein, N., Albar, L., Nublat, A. and Keller, B. (2003) Map-based isolation of the leaf rust disease resistance gene Lr10 from the hexaploid wheat (*Triticum aestivum* L.) genome. *Proc. Natl. Acad. Sci. USA* 100, 15253–15258.
- Fu, H. and Dooner, H.K. (2002) Intraspecific violation of genetic colinearity and its implications in maize. *Proc. Natl. Acad. Sci. USA* 9914, 9573–9578.
- Gaut, B.S. and Doebley, J.F. (1997) DNA sequence evidence for the segmental allotetraploid origin of maize. *Proc. Natl. Acad. Sci. USA* 9413, 6809–6814.
- Gaut, B.S. (2001). Patterns of chromosomal duplication in maize and their implications for comparative maps of the grasses. *Genome Res.* 11, 55–66.
- Gaut, B.S. (2002) Evolutionary dynamics of grass genomes. *New Phytol.* 154, 15–28.
- Goff, S.A., Ricke, D., Lan, T.H., Presting, G. et al. (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. japonica). *Science* 296, 92–100.

- Gottwald, S., Stein, N., Borner, A., Sasaki, T. and Graner, A. (2004) The gibberellic-acid insensitive dwarfing gene *sdw3* of barley is located on chromosome 2HS in a region that shows high colinearity with rice chromosome 7L. *Mol. Genet. Genomics* 271, 426–436.
- Griffiths, S., Dunford, R.P., Coupland, G. and Laurie, D.A. (2003) The evolution of CON-STANS-like gene families in barley, rice, and Arabidopsis. *Plant Physiol.* 131, 1855–1867.
- Griffiths, S., Sharp, R., Foote, T.N., Bertin, I., Wanous, M., Reader, S., Colas, I. and Moore, G. (2006) Molecular characterization of Ph1 as a major chromosome pairing locus in polyploid wheat. *Nature* 439, 749–752.
- Gu, Y.Q., Coleman-Derr, D., Kong, X. and Anderson, O.D. (2004) Rapid genome evolution revealed by comparative sequence analysis of orthologous regions from four triticeae genomes. *Plant Physiol.* 135, 459–470.
- Gu, Y.Q., Salse, J., Coleman-Derr, D., Dupin, A., Crossman, C., Lazo, G.R., Huo, N., Belcram, H., Ravel, C., Charret, G., Charles, M., Anderson, O.D. and Chalhou, B. (2006) Types and rates of sequence evolution at the high-molecular-weight glutenin locus in hexaploid wheat and its ancestral genomes. *Genetics* 1743, 1493–1504.
- Guyot, R., Yahiaoui, N., Feuillet, C. and Keller, B. (2004). In silico comparative analysis reveals a mosaic conservation of genes within a novel colinear region in wheat chromosome 1AS and rice chromosome 5S. *Funct. Integr. Genomics* 4, 47–58.
- Hampson, S., McLysaght, A., Gaut, B. and Baldi, P. (2003) LineUp: statistical detection of chromosomal homology with application to plant comparative genomics. *Genome. Res.* 13, 1–12.
- Hasterok, R., Marasek, A., Donnison, I.S., Armstead, I., Thomas, A., King, I.P., Wolny, E., Idziak, D., Draper, J. and Jenkins, G. (2006) Alignment of the genomes of brachypodium distachyon and temperate cereals and grasses using bacterial artificial chromosome landing with fluorescence in situ hybridization. *Genetics* 173, 349–362.
- Herrmann, R.G., Martin, R., Busch, W., Wanner, G. and Hohmann, U. (1996) Physical and topographical mapping among Triticeae chromosomes. *Symp. Soc. Exp. Biol.* 50, 25–30.
- Hohmann, U., Endo, T.R., Gill, K.S. and Gill, B.S. (1994) Comparison of genetic and physical maps of group 7 chromosomes from *Triticum aestivum* L. *Mol. Gen. Genet.* 2455, 644–653.
- Huang, S.X., Sirikhachornkit, A., Faris, J.D., Su, X.J., Gill, B.S., Haselkorn, R. and Gornicki, P. (2002) Phylogenetic analysis of the acetyl-CoA carboxylase and 3-phosphoglycerate kinase loci in wheat and other grasses. *Plant Mol. Biol.* 48, 805–820.
- Inada, D.C., Bashir, A., Lee, C., Thomas, B.C., Ko, C., Goff, S.A. and Freeling, M. (2003) Conserved noncoding sequences in the grasses. *Genome Res.* 13, 2030–2041.
- International Rice Genome Sequencing Project. (2005) The map-based sequence of the rice genome. *Nature* 436, 793–800.
- Isidore, E., Scherrer, B., Chaloub, B., Feuillet, C. and Keller, B. (2005) Ancient haplotypes resulting from extensive molecular rearrangements in the wheat A genome have been maintained in species of three different ploidy levels. *Genome Res.* 15, 526–536.
- Jaiswal, P., Ni, J., Yap, I., Ware, D. et al. (2006) Gramene: a bird's eye view of cereal genomes. *Nucl. Acids Res.* 34, 717–723.
- Janda, J., Bartos, J., Safar, J., Kubalaková, M. et al. (2004) Construction of a subgenomic BAC library specific for chromosomes 1D, 4D and 6D of hexaploid wheat. *Theor. Appl. Genet.* 109, 1337–1345.
- Janda, J., Šafář, J., Kubaláková, M., Bartoš, J. et al. (2006) Advanced resources for plant genomics: BAC library specific for the short arm of chromosome 1B. *Plant J.* 47, 977–986.
- Kaplin, N.J., Braun, D.M., Penterman, J., Goff, S.A. and Freeling, M. (2002) Utility and distribution of conserved noncoding sequences in the grasses. *Proc. Natl. Acad. Sci.* 99, 6147–6151.
- Kellogg, E.A. (2001) Evolutionary history of the grasses. *Plant Phys.* 125, 1198–1205.
- Kilian, A., Chen, J., Han, F., Steffenson, B. and Kleinhofs, A. (1997) Towards map-based cloning of the barley stem rust resistance genes *rpg1* and *rpg4* using rice as an intergenomic cloning vehicle. *Plant Mol. Biol.* 35, 187–195.

- Kishimoto, N., Higo, H., Abe, K., Arai, S., Saito, A. and Higo, K. (1994) Identification of the duplicated segments in rice chromosomes 1 and 5 by linkage analysis of cDNA markers of known functions. *Theor. Appl. Genet.* 88, 722–726.
- Koller, O.L. and Zeller, F.J. (1976) Homoeologous relationships of rye chromosomes 4R and 7R with wheat chromosomes. *Genet. Res.* 28, 177–188.
- Kubalakova, M., Vrana, J., Cihalikova, J., Simkova, H. and Dolezel, J. (2002) Flow karyotyping and chromosome sorting in bread wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* 104, 1362–1372.
- Lai, J., Li, Y., Messing, J. and Dooner, H.K. (2005) Gene movement by Helitron transposons contributes to the haplotype variability of maize. *Proc. Natl. Acad. Sci. USA* 10225, 9068–9073.
- La Rota, M. and Sorrells, M.E. (2004) Comparative DNA sequence analysis of mapped wheat ESTs reveals the complexity of genome relationships between rice and wheat. *Funct. Integr. Genomics* 4, 34–46.
- Leister, D., Kurth, J., Laurie, D.A., Yano, M., Sasaki, T., Devos, K., Graner, A. and Schulze-Lefert, P. (1998) Rapid reorganization of resistance gene homologues in cereal genomes. *Proc. Natl. Acad. Sci. USA* 95, 370–375.
- Li, W. and Gill, B.S. (2002) The colinearity of the Sh2/A1 orthologous region in rice sorghum and maize is interrupted and accompanied by genome expansion in the Triticeae. *Genetics* 160, 1153–1162.
- Lin, Y.R., Schertz, K.F. and Paterson, A.H. (1995) Comparative analysis of QTLs affecting plant height and maturity across the poaceae, in reference to an interspecific sorghum population. *Genetics* 141, 391–411.
- Lisch, D.R., Freeling, M., Langham, R.J. and Choy, M. (2001) Mutator transposase is widespread in the grasses. *Plant Physiol.* 1253, 1293–1303.
- Lisch, D. (2002) Mutator transposons. *Trends Plant Sci.* 711, 498–504.
- Liu, C.J., Atkinson, M.D., Chinoy, C.N., Devos, K.M. and Gale, M.D. (1992) Nonhomoeologous translocations between group 4, 5, and 7 chromosomes within wheat and rye. *Theor. Appl. Genet.* 83, 305–312.
- Lockton, S. and Gaut, B.S. (2005) Plant conserved non-coding sequences and paralogue evolution. *Trends Genet.* 21, 60–65.
- Moore, G., Devos, K.M., Wang, Z. and Gale, M.D. (1995) Cereal genome evolution: grasses, line up and form a circle. *Curr. Biol.* 5, 737–739.
- Morgante, M., Brunner, S., Pea, G., Fengler, K., Zuccolo, A. and Rafalski, A. (2005) Gene duplication and exon shuffling by helitron-like transposons generate intraspecies diversity in maize. *Nat. Genet.* 379, 997–1002.
- Nagamura, Y., Inoue, T., Antonio, B.A. et al. (1995) Conservation of duplicated segments between rice chromosomes 11 and 12. *Breed. Sci.* 45, 373–376.
- Naranjo, T., Roca, A., Goicoechea, P.G. and Giraldez, R. (1987) Arm homoeology of wheat and rye chromosomes. *Genome* 29, 873–882.
- Naranjo, T. and Fernandez-Rueda, P. (1991) Homoeology of rye chromosome arms to wheat. *Theor. Appl. Genet.* 82, 577–586.
- Park, Y.J., Dixit, A., Yoo, J.W. and Bennetzen, J. (2004) Further evidence of microcolinearity between barley and rice genomes at two orthologous regions. *Mol. Cells* 173, 492–502.
- Paterson, A.H., Lin, Y.R., Li, Z.K., Schertz, K.F., Doebley, J.F., Pinson, S.R.M., Liu, S.C., Stansel, J.W. and Irvine, J.E. (1995) Convergent domestication of cereal crops by independent mutations at corresponding genetic loci. *Science* 269, 1714–1718.
- Paterson, A.H., Bowers, J.E. and Chapman, B.A. (2004) Ancient polyploidization predating divergence of the cereals, and its consequences for comparative genomics. *Proc. Natl. Acad. Sci. USA* 101, 9903–9908.
- Paterson, A.H., Freeling, M. and Sasaki, T. (2005) Grains of knowledge: genomics of model cereals. *Genome Res.* 1512, 1643–1650.
- Peng, J.R., Richards, D.E., Hartley, N.M., Murphy, G.P. et al. (1999) ‘Green revolution’ genes encode mutant gibberellin response modulators. *Nature* 400, 256–261.

- Pourkheirandish, M., Wicker, T., Stein, N., Fujimura, T. and Komatsuda, T. (2007) Analysis of the barley chromosome 2 region containing the six-rowed spike gene *vrs1* reveals a breakdown of rice-barley micro collinearity by a transposition. *Theor. Appl. Genet.* 1148, 1357–1365.
- Qi, L.L., Echaliier, B., Chao, S., Lazo, G.R. et al. (2004) A chromosome bin map of 16,000 expressed sequence tag loci and distribution of genes among the three genomes of polyploid wheat. *Genetics* 168, 701–712.
- Ramakrishna, W., Dubcovsky, Y., Park, Y.J., Busso, C.S., Emberton, J., SanMiguel, P. and Bennetzen, J.L. (2002a) Different types and rates of genome evolution detected by comparative sequence analysis of orthologous segments from four cereal genomes. *Genetics* 162, 1389–1400.
- Ramakrishna, W., Emberton, J., Ogden, M., SanMiguel, P. and Bennetzen, J.L. (2002b) Structural analysis of the maize Rpl complex reveals numerous sites and unexpected mechanisms of local rearrangement. *Plant Cell* 14, 3213–3223.
- Safar, J., Bartos, J., Janda, J., Bellec, A. et al. (2004) Dissecting large and complex genomes: flow sorting and BAC cloning of individual chromosomes from bread wheat. *Plant J.* 39, 960–968.
- Salse, J., Piegu, B., Cooke, R. and Delseny, M. (2002) Synteny between *Arabidopsis thaliana* and rice at the genome level: a tool to identify conservation in the ongoing rice genome sequencing project. *Nucl. Acids Res.* 30, 2316–2328.
- Salse, J., Piegu, B., Cooke, R. and Delseny, M. (2004) New in silico insight into the synteny between rice (*Oryza sativa* L.) and maize (*Zea mays* L.) highlights reshuffling and identifies new duplications in the rice genome. *Plant J.* 38, 396–409.
- Salse, J. and Feuillet, C. (2007) Comparative genomics of cereals. In: R. Varshney and R. Tuberosa (Eds.), *Genomics-Assisted Crop Improvement* (Chapter 18). Springer Verlag, Dordrecht, The Netherlands, pp. 177–205.
- Salse, J., Bolot, S., Throude, M., Jouffe, V., Piegu, B., Masood, U., Calcagno, T., Cooke, R., Delseny, M. and Feuillet, C. (2008) Identification and characterization of conserved duplications between rice and wheat provide new insight into grass genome evolution. *Plant Cell* 20, 11–24.
- Sasaki, T., Matsumoto, T., Yamamoto, K., Sakata, K. et al. (2002) The genome sequence and structure of rice chromosome 1. *Nature* 420, 312–316.
- Scherrer, B., Isidore, E., Klein, P., Kim, J.S., Bellec, A., Chalhou, B., Keller, B. and Feuillet, C. (2005) Large intraspecific haplotype variability at the Rph7 locus results from rapid and recent divergence in the barley genome. *Plant Cell* 17, 361–374.
- Singh, N.K., Raghuvanshi, S., Srivastava, S.K., Gaur, A. et al. (2004) Sequence analysis of the long arm of rice chromosome 11 for rice-wheat synteny. *Funct. Integr. Genomics* 4, 102–117.
- Smith, D.B. and Flavell, R.B. (1975) Characterization of the wheat genome by renaturation kinetics. *Chromosoma* 50, 223–242.
- Song, R. and Messing, J. (2003) Gene expression of a gene family in maize based on noncollinear haplotypes. *Proc. Natl. Acad. Sci. USA* 10015, 9055–9060.
- Sorrells, M.E., LaRota, M., Bermudez-Kandianis, C.E., Greene, R.A. et al. (2003) Comparative DNA sequence analysis of wheat and rice genomes. *Genome Res.* 13, 1818–1827.
- Sorrells, M.E. (2004) Cereal Genomics research in the post-genomic era. In: P.K. Gupta and R.K. Varshney (Eds.), *Cereal Genomics*. Kluwer Academic Publishers, Dordrecht, pp. 559–584.
- Stein, N., Feuillet, C., Wicker, T., Schlagenhauf, E. and Keller, B. (2000) Subgenome chromosome walking in wheat: a 450-kb physical contig in *Triticum monococcum* L. spans the *Lr10* resistance locus in hexaploid wheat (*Triticum aestivum* L.). *Proc. Natl. Acad. Sci. USA* 97, 13436–13441.
- Stein, N. (2007) Triticeae genomics: advances in sequence analysis of large genome cereal crops. *Chromosome Res.* 151, 21–31.

- Stein, N., Prasad, M., Scholz, U., Thiel, T., Zhang, H., Wolf, M., Kota, R., Varshney, R.K., Perovic, D., Grosse, I. and Graner, A. (2007) A 1,000-loci transcript map of the barley genome: new anchoring points for integrative grass genomics. *Theor. Appl. Genet.* 114(5), 823–839.
- Suoniemi, A., Anamthawat-Jónsson, K., Arna, T. and Schulman, A.H. (1996a) Retrotransposon BARE-1 is a major, dispersed component of the barley (*Hordeum vulgare* L.) genome. *Plant Mol. Biol.* 30(6), 1321–1329.
- Suoniemi, A., Narvanto, A. and Schulman, A.H. (1996b) The BARE-1 retrotransposon is transcribed in barley from an LTR promoter active in transient assays. *Plant Mol. Biol.* 31(2), 295–306.
- Sutton, T., Whitford, R., Baumann, U., Dong, C., Able, J.A. and Langridge, P. (2003) The Ph2 pairing homoeologous locus of wheat (*Triticum aestivum*): identification of candidate meiotic genes using a comparative genetics approach. *Plant J.* 36(4), 443–456.
- Sutton, T., Baumann, U., Hayes, J., Collins, N.C., Shi, B.J., Schnurbusch, T., Hay, A., Mayo, G., Pallotta, M., Tester, M. and Langridge, P. (2007) Boron-toxicity tolerance in barley arising from efflux transporter amplification. *Science* 318(5855), 1446–1449.
- Swigonová, Z., Lai, J., Ma, J., Ramakrishna, W., Llaca, V., Bennetzen, J.L. and Messing, J. (2004) Close split of sorghum and maize genome progenitors. *Genome Res.* 14(10A), 1916–1923.
- The Rice Chromosome 10 Sequencing Consortium. (2003) In-depth view of structure, activity, and evolution of rice chromosome 10. *Science* 300, 1566–1569.
- The Rice Chromosome 3 Sequencing Consortium. (2005) Sequence, annotation, and analysis of synteny between rice chromosome 3 and diverged grass species. *Genome Res.* 15, 1284–1291.
- The Rice Full-Length cDNA Consortium. (2003) Collection, mapping, and annotation of over 28 000 cDNA clones from japonica rice. *Science* 301, 376–379.
- Turner, A., Beales, J., Faure, S., Dunford, R.P. and Laurie, D.A. (2005) The pseudo-response regulator Ppd-H1 provides adaptation to photoperiod in barley. *Science* 310, 1031–1034.
- Vandepoele, K., Saey, Y., Simillion, C., Raes, J. and Van de Peer, Y. (2002) The automatic detection of homologous regions (ADHoRe) and its application to microcolinearity between Arabidopsis and rice. *Genome Res.* 12, 1792–1801.
- Vandepoele, K., Simillion, C. and Van de Peer, Y. (2003) Evidence that rice and other cereals are ancient aneuploids. *Plant Cell* 15, 2192–2202.
- Van Deynze, A.E., Nelson, J.C., O'Donoghue, L.S., Ahn, S.N., Siripoonwiwat, W., Harrington, S.E., Yglesias, E.S., Braga, D.P., McCouch, S.R. and Sorrells, M.E. (1995) Comparative mapping in grasses. Oat relationships. *Mol. Gen. Genet.* 249(3), 349–356.
- Van Deynze, A.E., Nelson, J.C., Yglesias, E.S., Harrington, S.E., Braga, D.P., McCouch, S.R. and Sorrells, M.E. (1995, Oct) Comparative mapping in grasses. Wheat relationships. *Mol. Gen. Genet.* 25, 248(6), 744–754.
- Vogel, J.P., Gu, Y.Q., Twigg, P., Lazo, G.R., Laudencia-Chingcuanco, D., Hayden, D.M., Donze, T.J., Vivian, L.A., Stamova, B. and Coleman-Derr, D. (2006) EST sequencing and phylogenetic analysis of the model grass *Brachypodium distachyon*. *Theor. Appl. Genet.* 113(2), 186–195.
- Wang, X., Shi, X., Hao, B., Ge, S. and Luo, J. (2005). Duplication and DNA segmental loss in the rice genome: implications for diploidization. *New Phytol.* 165, 937–946.
- Waugh, R., McLean, K., Flavell, A.J., Pearce, S.R., Kumar, A., Thomas, B.B. and Powell, W. (1997) Genetic distribution of Bare-1-like retrotransposable elements in the barley genome revealed by sequence-specific amplification polymorphisms (S-SAP). *Mol. Gen. Genet.* 253(6), 687–694.
- Wei, F., Coe, E., Nelson, W., Bharti, A.K., Engler, F., Butler, E., Kim, H., Goicoechea, J.L., Chen, M., Lee, S., Fuks, G., Sanchez-Villeda, H., Schroeder, S., Fang, Z., McMullen, M., Davis, G., Bowers, J.E., Paterson, A.H., Schaeffer, M., Gardiner, J., Cone, K., Messing, J., Soderlund, C. and Wing, R.A. (2007) Physical and genetic structure of the maize genome reflects its complex evolutionary history. *PLoS Genet.* 3(7), e123.

- Wendel, J.F., Stuber, C.W., Goodman, M.M. and Beckett, J.B. (1989) Duplicated plastid and triplicated cytosolic isozymes of triosephosphate isomerase in maize (*Zea mays* L.). *J. Hered.* 3, 218–228.
- Wicker, T., Yahiaoui, N., Guyot, R., Schlagenhauf, E., Liu, Z.D., Dubcovski, J. and Keller, B. (2003) Rapid genome divergence at orthologous low molecular weight glutenin loci of the A and Am genomes of wheat. *Plant Cell* 15, 1186–1197.
- Wicker, T., Yahiaoui, N. and Keller, B. (2007) Contrasting rates of evolution in pm3 Loci from three wheat species and rice. *Genetics* 177(2), 1207–1216.
- Yahiaoui, N., Srichumpa, P., Dudler, R. and Keller, B. (2004) Genome analysis at different ploidy levels allows cloning of the powdery mildew resistance gene Pm3b from hexaploid wheat. *Plant J.* 37, 528–538.
- Yan, L., Loukoianov, A., Tranquilli, G., Helguera, M., Fahima, T. and Dubcovsky, J. (2003) Positional cloning of the wheat vernalization gene VRN1. *Proc. Natl. Acad. Sci. USA* 100, 6263–6268.
- Yan, L., Loukoianov, A., Blechl, A., Tranquilli, G., Ramakrishna, W., SanMiguel, P., Bennetzen, J.L., Echenique, V. and Dubcovsky, J. (2004) The wheat VRN2 gene is a flowering repressor down-regulated by vernalization. *Science* 303, 1640–1644.
- Yan, L., Fu, D., Li, C., Blechl, A., Tranquilli, G., Bonafede, M., Sanchez, A., Valarik, M., Yasuda, S. and Dubcovsky, J. (2006) The wheat and barley vernalization gene VRN3 is an orthologue of FT. *Proc. Natl. Acad. Sci. USA* 103(51), 19581–19586.
- Yu, J., Hu, S., Wang, J., Wong, G.K.S. et al. (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). *Science* 296, 79–92.
- Yu, J., Wang, J., Lin, W., Li, S. et al. (2005) The genomes of *Oryza sativa*: a history of duplications. *PLoS Biol.* 3, 266–281.

Chapter 18

Genomics of Tolerance to Abiotic Stress in the Triticeae

Marco Maccaferri, Maria Corinna Sanguineti, Silvia Giuliani
and Roberto Tuberosa

Abstract Genomics platforms offer unprecedented opportunities to identify, select and in some cases clone the genes and the quantitative trait loci (QTLs) that govern the tolerance of Triticeae to abiotic stresses and, consequently, grain yield. Transcriptome profiling and the other “omics” platforms provide further information to unravel gene functions and validate the role of candidate genes. This review provides a synopsis of the main results on the studies that have investigated the genomics of Triticeae crops under conditions of abiotic constraints. With their rich biodiversity and high functional plasticity in response to environmental stresses, Triticeae crops provide an ideal ground for taking full advantage of the opportunities offered by genomics approaches. Ultimately, the practical impact of the knowledge and materials generated through genomics-based approaches will depend on their integration and exploitation within the extant breeding programs.

18.1 Introduction

The recent climate changes brought about by global warming have increased the frequency and severity of crop losses due to abiotic constraints. Triticeae crops are prevalingly grown under rainfed conditions in marginal areas and are thus more exposed to the negative effects of drought, low nutrients and toxic minerals, often present simultaneously (Collins et al. 2008). Because wheat and barley are the staple food for billions of people and their demand is expected to soar to meet the needs of a burgeoning population, improving tolerance to abiotic stress will become an increasingly important priority for most breeding programs. Although conventional breeding has allowed for a steady increase in yield under environmentally constrained conditions (Blum 1988; Ceccarelli and

R. Tuberosa (✉)

Department of Agroenvironmental Sciences and Technology, University of Bologna,
Bologna, Italy

e-mail: roberto.tuberosa@unibo.it

Grando 1996; Borlaug and Dowswell 2005; Reynolds et al. 2005; Borlaug 2007; Li et al. 2007; Munns and Richards 2007; Richards et al. 2007), our capacity to dissect the genetic basis of abiotic stress tolerance has been greatly enhanced by the introduction of the genomics platforms that allow us to identify the genes and quantitative trait loci (QTLs) governing genetic variation at relevant loci (Tanksley 1993; Ribaut and Hoisington 1998; Tuberosa et al. 2002a; Munns 2005; Varshney et al. 2005; Varshney and Tuberosa 2007). Starting from the early 1990s, an increasing number of studies have reported on genes and QTLs in plants exposed to abiotic stress (Blum 1996; Zhu et al. 1997; Hasegawa et al. 2000; Cattivelli et al. 2002, 2008; Tuberosa et al. 2002c; Bohnert et al. 2006; Tuberosa and Salvi 2006; Vij and Tyagi 2007; see also <http://www.plantstress.com>). New dimensions for deciphering the genetic control of the response to abiotic stress have been ushered in by comparative mapping (Salse and Feuillet 2007), sequencing (Goff et al. 2002) and post-genomics platforms (Hazen and Kay 2003; Hirai et al. 2005; Alonso and Ecker 2006).

This review compiles the main results of studies that have applied genomics approaches to dissect the genetic basis of tolerance to abiotic stress in the Triticeae. In view of the prevailing quantitative nature of the inheritance of abiotic stress tolerance, particularly in terms of yield, we have focused primarily, but not exclusively, on the QTLs that have been shown to influence the adaptive response to environmental constraints. Owing to the vast number of studies that have dealt with the topic of this survey, for the sake of conciseness it has not been possible to cover each report in much detail. Nonetheless, Tables 18.1A, 18.1B and 18.2 summarise the essential features of the majority of relevant studies on QTLs, single genes and microarrays, respectively, and provide a glimpse into the functional complexity that underlines the performance of the Triticeae crops under conditions of abiotic stress.

18.2 Searching QTLs and Genes for Tolerance to Abiotic Stress

Several reviews have described methods and approaches required to identify and clone QTLs (Tanksley 1993; Lee 1995; Doerge 2002, 2007; Tuberosa et al. 2002b; Salvi and Tuberosa 2005, 2007). Most of the QTL results reported to date have been obtained through linkage mapping of populations derived from biparental crosses. More recently, association mapping using panels of germ-plasm accessions has provided an additional means to identify genes and/or QTLs for target traits (Gupta et al. 2005; Burke et al. 2007; Veyrieras et al. 2007). Although most of the association mapping studies published so far have targeted traits (e.g. resistance to biotic stress and quality traits) with a genetic basis less complex than abiotic stress tolerance, applications of the association approach to the latter traits have been reported in barley (Pakniyat et al. 1997) and durum wheat (Sanguineti et al. 2007). A valuable feature of association mapping is the possibility of resolving QTLs without the time-consuming

Table 18.1A Main features of a number of studies targeting the identification of QTLs for tolerance to abiotic stresses in the Triticeae

Genetic materials and references	Traits	Chromosomes with major QTLs	Other QTLs and comments
Drought – Barley			
Tadmor × (Er/Apm): 187 RILs evaluated in growth chamber under WW and WS.	RWC, no. of leaves of main tiller, TN, shoot biomass.	1 QTL for no. of leaves on 1H.	WW and WS: 100 and 14% of field capacity. WS: 2 QTLs for RWC, 4 QTLs for no. of leaves; WW: 1 QTL for RWC; 1 QTL for no. leaves. Epistatic interactions (in WS only) for RWC and shoot biomass.
Teulat et al. (1997).			
Tadmor × (Er/Apm): 187 RILs evaluated in growth chamber under WW and WS.	RWC, OP, OPFT, OA at early growth stage.	No major QTL.	WS vs. WW: 1 QTL for OA; WW: 1 QTL for RWC, 1 QTL for OPFT; WS: 3 QTLs for RWC, 4 QTLs for OP and 2 QTLs for OPFT, co-location of QTLs for RWC, OP, OA on 7H and 6H. Epistatic effects in both WW and WS.
Teulat et al. (1998).			
Tadmor × (Er/Apm): 167 RILs evaluated in growth chamber under WW and WS.	RWC, OP, OPFT, WSC, WSC at full turgor.	No major QTL.	Most QTLs of Teulat et al. (1998) confirmed + new ones. Total of 32 QTLs in 13 regions + epistasis. WW: no QTL for WSC and 1 for WSC at full turgor; WS: 1 QTL for WSC.
Teulat et al. (2001a).			
Tadmor × (Er/Apm): 167 RILs evaluated in 6 Mediterranean envs. (rainfed and irrigated).	GY, GY components, HD, PH.	1 QTL for PH on 3H and 1 QTL on 6H.	Major QTLs common to 4–6 envs. QTL clusters in several envs.; most QTLs common to irrigated and rainfed envs. and others specific. 24 consistent QTLs: 11 QTLs only additive effects, 7 QTLxE and 6 additive + QTLxE.
Teulat et al. (2001b).			
Tadmor × (Er/Apm): 167 RILs evaluated in 3 Mediterranean envs. (rainfed and irrigated).	CID.	No major QTL.	10 QTLs: 1 QTL specific for 1 env., 2 QTLxE and 7 additive + QTLxE.
Teulat et al. (2002).			Co-location of 8 QTLs for CDI with QTLs for water status (measured in Teulat et al. 2001a), OA, agronomic traits (Teulat et al. 2001b).

Table 18.1A (continued)

Genetic materials and references	Traits	Chromosomes with major QTLs	Other QTLs and comments
Tadmor × (Er/Apm): 167 RILs evaluated in 3 Mediterranean envs. (rainfed and irrigated), Teulat et al. (2003).	RWC.	1 on 6HL.	9 QTLs, 4 of which across different envs. and 1 on 6HL that could be of interest for marker-assisted selection to improve tolerance to drought.
Arta × <i>H. spontaneum</i> : 194 RILs evaluated in 2 Mediterranean locations for 2 years. Baum et al. (2003).	HD, PH, SPAD, GY, BY, TN, KW, GPC, etc.	1 QTL for HD on 7H and 1 QTL for PH on 3H.	10 QTLs for BY (1 in more envs. on 3H); 6 QTLs for GY (3 env.-specific and 3 on 1H and 3H in more envs.; positive allele from barley); 21 QTLs for KW; 8 QTLs for HD; 3 QTLs for GPC (favourable allele from <i>H. spontaneum</i> ; 2 QTLs co-located with QTLs for KW with plus allele from barley).
Barke × <i>Hordeum spontaneum</i> : 123 DHs from BC ₁ F ₂ plants evaluated in 3 envs. (Italy, Morocco, Tunisia) under rainfed conditions. Talamè et al. (2004).	GrH, HD, PH, PedL, PH, KW, GY.	No major QTL.	Total of 80 QTLs (from 3 QTLs for GrH to 18 QTLs for HD). 11–16 QTLs for GY per env. (10 consistent across envs.). <i>H. spontaneum</i> contributed the allele increasing GY at 6 QTLs, 2 of which (chr. 2H, <i>Bmac0093</i> and chr. 5H, <i>Bmac0684</i>) common to the 3 envs. and with the highest additive effect.
Steptoe × Morex: 72 DHs evaluated in the field under WS conditions in Iran for 2 years. Peighambari et al. (2005).	HD, PH, KW, GY, PrC, SL, etc.	1 QTL for PH and SL on 3H, R ² > 20%.	23 QTLs on all chrs. except 6H. 1 QTL for GY, 3 QTLs for KW, 3 QTLs for HD, 3 QTLs for PH, 2 QTLs for SL, 2 QTLs for PrC etc. QTL for PH on 3H already reported as stable across environments by Hayes and Iyambo (1994).
Drought – Wheat Dharwar Dry/Sitta: 140 F ₆ evaluated in Mexico under WS and WW for 2 years. Kirigwi et al. (2007).	DSI and associated traits, BY rate, HD, etc.	1 QTL on 4A for DSI.	Minor QTLs were not detected because the molecular analysis was limited to the major QTL found in a preliminary molecular screening.

Table 18.1A (continued)

Genetic materials and references	Traits	Chromosomes with major QTLs	Other QTLs and comments
W7984/Opata 85: 114 RILs in the field for 2 years. Desiccation treatment after anthesis. Salem et al. (2007).	Post-anthesis drought TI (stem reserve mobilization).	3 QTLs on 2DS, 5DS and 7DL for drought tolerance.	2DS-QTL proximal: $R^2 = 42\%$; 5DS-QTL centromeric: $R^2 = 37\%$; 7DL-QTL distal: $R^2 = 21\%$. The 7DL-QTL coincident with a QTL for osmotic stress tolerance at seedling stage (Landjeva et al. 2008).
W7984/Opata 85: 114 RILs evaluated at seedling stage under stress (12% PEG) and control conditions. Landjeva et al. (2008).	RL, shoot length, root/shoot ratio, etc. Various TI.	1 cluster (5 QTLs) on 1AL for RL, etc.; 1 QTL on 2DS for TI.	Minor QTLs on 6DL and 7DL for RL, root/shoot length under osmotic stress. Constitutive QTLs on several chr. arms (1AS, 1BL, 2DS, 5BL, 6BL) both under stressed and control conditions.
Jennah Khetifa/Cham 1: 110 RILs evaluated under controlled conditions. Kubo et al. (2007).	Rno penetrating the PV disc, total Rno, RPA, etc.	1 QTL on 6A for RPA and Rno penetrating the PV disc.	The favourable allele was inherited from Jennah Khetifa.
Hanxuan 10/Lumai 14: 150 RILs in the field under rainfed and irrigated conditions. Yang D.L. et al. (2007a).	Stem-WSC and GW measured from flowering to maturity.	No major QTL.	Complex genetic control: chrs. 1A, 1D, 2D, 4A, 6B, 7B, 7D. 48 additive QTLs; 62 pairs of epistatic QTLs; R^2 ranging from 1 up to 10%
Hanxuan 10/Lumai 14: 150 RILs evaluated in the field under rainfed and irrigated conditions. Yang D.L. et al. (2007b).	CFKP (Fv, Fm, Fv/Fm, Fv/Fo) and ChlC.	No major QTL.	CFKP: QTLs on 1B, 3B, 4D, 6A, 7A, 7D; WS: 11 QTLs and WW: 7 QTLs (R^2 from 8 to 72%). QTL \times water availability interaction. ChlC: QTLs on 1A, 5A, 7A.

Table 18.1A (continued)

Genetic materials and references	Traits	Chromosomes with major QTLs	Other QTLs and comments
Beaver/Soissons: 48–65 DHs evaluated in the field at various levels of water availability. Verma et al. (2004); Snape et al. (2007). Spark/Rialto: 144 DHs evaluated in the field at various levels of water availability. Snape et al. (2007). Chinese Spring substitution lines; Chinese Spring/SQ1: 139 F ₂ and 96 DHs evaluated under controlled conditions. Quarrie et al. (1994a,b). Chinese Spring/SQ1: 96 DHs evaluated in 24 envs. with a 10-fold range yield. Quarrie et al. (2005). Extreme temperature –	FLS and GY. Stem WSC. ABA in detached and partially dehydrated leaves (DLT). GY and GY components.	1 QTL on 2BL and 1 on 2DL. 1 QTL on the 1BL/1RS translocation. 1 QTL on 5AL.	FLS not influenced by HD and positively correlated with yield under variable environmental regimes. Major QTLs common to both water regimes. Major QTL: positive allele associated to 1R. Other QTLs: on 2B (<i>Xwmc332</i>), 3D (<i>Xgwm389</i>), 4B (<i>Xgwm251</i>) and 7A (<i>psp3094</i>). Complex genetic control of ABA under DLT. 5AL-QTL near to <i>Xpsr575</i> and <i>Xpsr426</i> , about 8 cM from <i>Xpsr426</i> . 7AL- and 7BL-QTLs in homeologous positions (distal region). QTL clusters detected in a range of envs.: 1AS, 1BL, 2BS, 3DL, 4AL, 4AS, 4BL, 4DL, 5AL, 5BS, 5DS, 5DL, 6BL, 7AL, 7BS, 7BL.
Barley Barley. Dicktoo × Morex: 100 DHs evaluated under controlled conditions and in the field. Hayes et al. (18.693); Pan et al. (1994).	LTT (controlled conditions); WSu (field).	1 stable QTL for WSu on 7H.	QTL clusters for growth habit, HD and photoperiod sensitivity mapped to the same regions.

Table 18.1A (continued)

Genetic materials and references	Traits	Chromosomes with major QTLs	Other QTLs and comments
Barley. Artia \times <i>Hordeum spontaneum</i> : 194 RILs evaluated in 2 field locations f or 2 years. Baum et al. (2003).	Cold damage (score).	No major QTL.	8 QTLs for cold damage; tolerant allele from barley at 5 QTLs, mainly on chr. 5H, and from <i>H. spontaneum</i> at 3 QTLs. <i>H. spontaneum</i> alleles showed a limited effect.
Barley. Nure/Tremois: 136 DHs under controlled conditions. Francia et al. (2004).	WSu, FT, VR, COR 14b and TMC-Ap3.	2 QTL clusters 20 cM apart on 5H.	Cluster 1: VR, WSu, FT, Fv/Fm; coincident with <i>Vrn-H1</i> , <i>aba2</i> , <i>dhn1</i> , <i>dhn2</i> . Cluster 2: WSu, FT, COR 14b, TMC-Ap3.
Barley: Arapiles/Franklin (A/F) 225 DHs; Galleon/Haruna Nijo (G/HN) 112 DHs; Amagi Nijo/WI2585 (AN/WI) 139 DHs. Spring radiation frost in the field (2 envs.). Reinheimer et al. (2004).	FIS, FIGD, TFID. Frost events between -3 and -4°C.	A/F: QTLs on 2HS and 5HL; G/HN: QTLs on 2H and 5H; AN/WI: QTLs on 2H and 5H.	A/F: 2HS-QTLs for FIS and FIGD (R^2 19–20%) coincident with <i>eps2</i> , an earliness per se gene. 5HL-QTLs for FIGD and TFID (R^2 18–32%). G/HN: 2H-QTL (R^2 42%, common to AN/WI) and 5H-QTL (R^2 44%) for FIS. AN/WI: QTLs for FIS (R^2 from 18 up to 32%).
Barley. Dicktoo/Morex: 136 DHs evaluated under controlled conditions. Skinner et al. (2006).	LTT under freezing conditions.	1 QTL on 5HL.	Genetic and physical mapping of <i>HvCBF</i> gene clusters and their regulators: the map position of <i>HvCBF</i> clusters did not coincide with <i>Fr-H1</i> . Additional LT tolerance QTLs on chrs. 1HL, 4HS, 4HL.
Barley. Nure/Tremois: 1849 F ₂ and selected F ₃ /F ₄ evaluated under controlled conditions. Francia et al. (2007).	FT, Fv/Fm at seedling stage.	2 QTLs for LTT on 5HL.	High resolution mapping of the <i>HvCBF</i> gene cluster at the <i>Fr-H2</i> mapping interval. <i>Fr-H2</i> fine mapped in a 4.6 cM interval.

Table 18.1A (continued)

Genetic materials and references	Traits	Chromosomes with major QTLs	Other QTLs and comments
Barley. Nure (N), Tremois (T) and NT recombinants in vernalized, non-vernalized and different photoperiod conditions. Stockinger et al. (2007).	Transcription level of candidate genes mapping in the QTL regions.	No major QTL.	Epistasis between <i>VrN-H1</i> / <i>Fr-H1</i> and <i>Fr-H2</i> (harboring the CBF gene cluster). Higher level of transcription of <i>Fr-H2</i> CBF in presence of <i>vrn-H1</i> winter allele, as compared to the spring allele. The CBF transcription level is increased at short days rather than at long days.
Extreme temperature – Wheat			
Hexaploid wheat. CS (<i>T. spelta</i> 5A)/CS (Ch 5A): RILs; CS/CS (Ch 5D): RILs. Pan et al. (1994); Laurie et al. (1995); Sarma et al. (1998); Snape et al. (2001).	FT under controlled conditions.	2 QTLs on 5A and 5D.	<i>Fr-5A1</i> and <i>Vrn-A1</i> are two distinct, closely linked loci. <i>Vrn</i> and <i>Fr</i> genes of homoeologous group 5 show conserved synteny with barley 5H, rye 5R and oat linkage group 24.
Hexaploid wheat. CS (<i>T. spelta</i> 5A)/CS (Ch 5A): 8 RL in controlled conditions. Galiba et al. (1997).	WSC in seedlings after cold hardening.	1 QTL on 5A.	Close linkage among <i>Fr-5A1</i> , <i>Vrn-A1</i> and QTLs for cold-induced WSC.
Hexaploid wheat. CS deletion lines evaluated under controlled conditions.	FT after cold hardening, HD without vernalization.	1 QTL on 5A.	Confirmation of linkage between the 2 genes (no pleiotropic effect of a single gene). The comparison between the physical and genetic maps for <i>Vrn-A1</i> and <i>Fr1</i> shows that the linear order is identical.
Sutka et al. (1999). Hexaploid wheat. CS (Ch 5B)/CS evaluated under controlled conditions. Toth et al. (2003).	FT.	1 QTL on 5BL.	<i>Fr-B1</i> (R^2 31%) mapped 40 cM from the centromere.

Table 18.1A (continued)

Genetic materials and references	Traits	Chromosomes with major QTLs	Other QTLs and comments
<i>T. monococcum</i> DV92/ <i>T. monococcum</i> ssp. <i>Aegilopoides</i> G3116. 74 RILs under controlled conditions (2 experiments). Vagujfalvi et al. (2003). <i>T. monococcum</i> DV92/ <i>T. monococcum</i> ssp. <i>Aegilopoides</i> G3116. 300 F ₂ plants. Miller et al. (2006).	FT (survival capability).	1 QTL on 5AL.	<i>Fr-A^m2</i> (R ² 40–48%) mapped 40 cM from the centromere and 30 cM proximal to <i>Fr-A1</i> . 1 e-QTL for the cold induced gene <i>Cor14b</i> co-mapped with <i>Fr-A^m2</i> . <i>Cbf3</i> was proposed as a candidate gene for <i>Fr-A^m2</i>
Hexaploid wheat. Ventnor/ Karl 92: 166 F ₂ F ₃ evaluated in the field. Yang et al. (2002).	FT.	1 QTL on 5AL.	Characterisation of the CBF family to identify <i>CBF</i> genes responsible for frost tolerance. Mapping of candidate genes in the QTL region: a cluster of 11 <i>CBF</i> genes is positioned at <i>Fr-A^m2</i> .
Hexaploid wheat. Renan/ Recital: 194 F ₇ RILs evaluated in 6 locations in France.	GF duration in presence of high temperatures.	2 QTLs on 1B and 5A.	R ² of the two major QTLs 11 and 12%, respectively. Tolerant types have a longer grain-filling period under stress (Ventnor is heat-tolerant).
Hexaploid wheat. Groos et al. (2003).	GPC, KW, GY + climatic factors.	1 QTL for GY on 7D.	GY: 8 QTLs (1 stable QTL, i.e. significant in more than 4 envs., on 7D; R ² up to 15%); GPC: 10 QTLs (4 stable on 2A, 3A, 4D, 7D); KW: 9 QTLs (3 stable on 2B, 5B, 7A). QTLs for differential reaction to climatic factors.
Hexaploid wheat. NILs at the <i>Vrn-A1</i> locus in 2 backgrounds. Controlled conditions: hydroponics. Limin and Fowler (2006).	LTT (acclimation, freezing test, recovery).	<i>VRN-A1</i> locus on 1 5A.	At <i>VRN-A1</i> locus, the <i>vrn-A1</i> winter allele confers greater LT tolerance compared to <i>Vrn-A1</i> (spring allele), irrespectively of the closely linked <i>Fr-A2</i> .

Table 18.1A (continued)

Genetic materials and references	Traits	Chromosomes with major QTLs	Other QTLs and comments
Hexaploid wheat. Norstar (N)/Winter Manitou (WM): 145 DHs; N/Cappelle Desprez: 256 DHs, N/Manitou: 152 DHs in controlled conditions. Båga et al. (2007).	LTT (acclimation, freezing test, recovery).	N/WM population: 1 QTL on 5A.	Cross between winter parents with different LTT to detect QTLs for LTT in addition to <i>vrn-1</i> . N/WM: 2 QTLs in total with the major QTL (chr. 5A, R^2 40%), 46 cM proximal to the <i>vrn-A1</i> locus and coincident with <i>Fr-A2</i> (<i>CBF14</i> and <i>CBF15</i>); second QTL (R^2 8%) on chr. 1D (<i>barc169</i>).
Salinity – Barley and wheat			
Barley. Steptoe/Morex (S/M) 149 DHs; Harrington/TR306 (H/T): 146 DHs evaluated under controlled conditions. Mano and Takeda (1997).	ST at germination, ST at seedling stage (3 experiments).	1 QTL for ST at germination on 5H in both pops.	Major 5H-QTL: R^2 41–47%. Different QTLs for ST at germination and seedling stage. S/M: 3 QTLs (4H, 6H, 5H) for ST at germination and 9 at seedling; H/T: 2 QTLs (1H, 5H) for ST at germination and 1 (5H) at seedling. > 5H-QTL coincides with an ABA-response QTL at germination.
Barley: Derkado/B83-12-21-5 156 DHs evaluated in: hydroponics. Ellis et al. (2002).	7 seedling traits: seedling growth, etc.	No major QTL.	A total of 12 QTLs for all seedling traits. 7 QTLs co-located with the dwarfing genes <i>sdw1</i> on chr. 3H and <i>ari-e.GP</i> on chr. 5H. QTLs for leaf growth (chrs. 2H and 3H) independent of dwarfing genes.
Wheat- <i>L. elongatum</i> amphiploid, disomic substitution lines in the field at 2 salinity levels + control. Omielan et al. (1991).	GY, BY, ion accumulation in leaves, ion transport etc.	No major QTL.	Increased salt tolerance due to Na^+ and Cl^- exclusion, K^+ inclusion and K^+ retranslocation. Chr. 3E of <i>L. elongatum</i> contributed to a 50% reduction of Na^+ in leaves. Several wheat chrs. controlled the ion exclusion/inclusion with the most notable effects on chr. 4D.

Table 18.1A (continued)

Genetic materials and references	Traits	Chromosomes with major QTLs	Other QTLs and comments
Wheat. Langdon disomic substitution lines D genome (controlled env.). Dubcovsky et al. (1996); Gorham et al. (1997).	Ion accumulation in leaves, K^+ / Na^+ discrimination.	<i>Kna1</i> gene on 4DL.	Greater discrimination (in favour of K^+ and against Na^+ accumulation) controlled by chr. 4DL: <i>Kna1</i> (distal third) is involved. Recombined <i>Kna1</i> 4B/4D genetic materials were obtained and <i>Kna1</i> was mapped within a short 2 cM interval.
Durum wheat. Line 149/Tamaroi: 100 F ₂ and 100 F _{2.3} + other populations under controlled conditions.	[Na^+] in the leaf measured 10 d after emergence.	1 QTL on 2AL.	2AL-QTL: R ² 35%. Line 149: source of the " Na^+ exclusion" trait for durum wheat (salt-resistance). <i>Nax1</i> reduces Na^+ concentration in leaves by retaining it in the sheaths.
Lindsay et al. (2004). Durum wheat. NILs for <i>Nax1</i> and/or <i>Nax2</i> under controlled conditions. James et al. (2006).	Na^+ and K^+ transport from roots to shoots.	<i>Nax1</i> on 2AL.	Lines with <i>Nax1</i> or <i>Nax2</i> have lower Na^+ in leaves, lower rates of Na^+ transport (not due to lower rates of net uptake from the soil) and higher rates of K^+ transport.
Hexaploid wheat. Opata 85/W7984; 114 RILs under controlled conditions. Ma et al. (2007).	Growth response at the germination and seedling stage.	No major QTL.	10 and 37 QTLs at the germination and seedling stages, respectively. QTL clusters on chromosomes 3AS, 3BL, 4DL, 6DL. QTLs on 3AS and 3BL effective at germination and seedling stages.
Toxic ions – Aluminium tolerance – Wheat and rye			
Hexaploid wheat. BH 1146/Anahuac: 97 RILs + 4DL deletion lines under controlled conditions (hydroponics). Milla and Gustafson (2001).	Root elongation.	<i>AltBH</i> on 4DL.	Genetic mapping on RILs + physical mapping on the distal region of the chromosome 4DL using wheat deletion lines for chromosome arm 4DL.

Table 18.1A (continued)

Genetic materials and references	Traits	Chromosomes with major QTLs	Other QTLs and comments
Hexaploid wheat. 5 DH populations under controlled conditions (hydroponics). Raman et al. (2005).	Root growth and/or haematoxylin staining.	<i>ALMT1</i> on 4D.	<i>ALMT1</i> cosegregated with Al tolerance. In 2 pops Al-tolerance cosegregates with greater capacity for efflux of malate from root apices. Genomic structure of the <i>ALMT1</i> gene.
Hexaploid wheat. Atlas66/Chisholm: 192 RILs under controlled conditions (hydroponics). Zhou et al. (2007).	Relative root growth.	1 QTL on 4DL.	The 4DL QTL (R^2 43–49%) co-segregated with the Al-activated malate transporter gene (<i>ALMT1</i>). Another minor QTL located on chr. 3BL.
Rye and wheat-rye addition lines under controlled conditions (hydroponics). Aniol and Gustafson (1984); Hede et al. (2001); Aniol (2004).	Root elongation after a short Al treatment.	3 QTLs on 6RS, 3R, 4RL.	Physical mapping in rye (<i>Secale cereale</i> L.) using wheat-rye addition lines.
Rye. M39A-1-6/M77A-1 F ₂ , F ₃ under controlled conditions (hydroponics). Miftahudin et al. (2005).	Root growth.	<i>Alt3</i> on 4RL.	Fine mapping of <i>Alt3</i> with RFLPs and initial cloning of a full-length candidate gene in rye. Further attempts to clone a full-length rye <i>Alt3</i> candidate gene will necessitate the creation of a rye large-insert library.
Rye. Ailes/Riodeva: three F ₂ populations (ca. 120–150 plants) under controlled conditions.	Root length (root re-growth in hydroponics).	<i>Alt3</i> on 4RL. <i>Alt4</i> on 7RS.	Mapping of <i>Alt3</i> , homolog of <i>Alp</i> gene in barley and <i>AltBH</i> in wheat (4DL). Mapping of EST-SSRs in rye (SCM), isolated from Triticeae ESTs involved in putative Al tolerance mechanisms in chrs. 3R, 4R, 5R. Mapping of the candidate gene <i>ScALMT1</i> , homologs of <i>TaALMT1</i> .
Matos et al. (2005); Fontecha et al. (2007); Matos et al. (2007).			

Table 18.1A (continued)

Genetic materials and references	Traits	Chromosomes with major QTLs	Other QTLs and comments
Toxic ions – Boron tolerance – Barley and wheat			
Barley. Clipper/Sahara377: DHs under controlled conditions (2 Boron levels). Jeffries et al. (1999); Karakousis et al. (2003); Sutton et al. (2007). Hexaploid wheat. Halberd/Cranbrook: 161 DHs evaluated in filter- and soil-based assays. Jeffries et al. (2000); Schnurbusch et al. (2007).	Leaf symptom expression, boron accumulation, RL, etc. Relative RL, leaf symptom, total DW, shoot B concentration.	1 QTL on 4H for most traits under high boron level. 1 QTL (<i>Bo1</i>) on 7BL for shoot B concentration.	Boron-tolerance major QTL on 4H; others on 2H (<i>Xcd0370</i> , <i>Bmac93</i>) and 3H (<i>AWBMA15</i> , <i>Bmac67</i>). Fine mapping of 4H QTL (6,720 meioses): the gene mapped within 0.15 cM between <i>XBM178</i> and <i>XBM162</i> . 2 QTLs on 7B and 7D (<i>Xpsr160</i>). <i>Bo1</i> confers up to 11% yield advantage in B toxic sites and a 74% reduction in shoot B concentration. Using wheat-rice synteny <i>Bo1</i> interval reduced to 1.8 cM; a co-dominant PCR marker (<i>AWW5L7</i>) co-segregates with <i>Bo1</i> .
Low nutrients – Barley and wheat			
Barley. Karl/Lewis 146 RILs evaluated in the field (2 years). Mickelson et al. (2003).	GY, GPC, N uptake, storage, remobilization etc.	1 QTL on 6H for GPC (R ² 46%)	No overlap between major GPC-QTL and QTLs for N remobilization. 2–8 QTLs for each N-related trait (R ² 8–20%) with major QTL clusters on 3H and 6H; at these clusters inefficient N remobilization associated with low GY.
Barley. Karl/Lewis 146 RILs evaluated in the field (2 years). Yang et al. (2004).	Leaf-N, amino-, carboxy-peptid-se activity, GY, etc.	No major QTL.	Comparison with Mickelson et al. (2003). QTLs for peptidase activities overlapped with leaf-N (3H), N remobilization (5H), leaf amino acids and leaf senescence (2H, 4H): 1 carboxy-peptidase enzyme involved in N remobilization from senescing leaves.

Table 18.1A (continued)

Genetic materials and references	Traits	Chromosomes with major QTLs	Other QTLs and comments
Hexaploid wheat. Renan/Recital: 194 RILs evaluated in the field (1 year, 6 envs.) Groos et al. (2003).	GPC, KW, GY.	1 for GY (R^2 16%) and GPC (R^2 10%) on 7D.	Stable QTLs (detected in at least 4 locations) for GPC on 2AS (<i>Xgwm400-Xgwm636</i>), 3A (<i>Xcfd79</i>), 4D (<i>Xcfd71</i>) and 7D. QTLs for KW on 2B, 5B and 7A. No negative correlation between GY and GPC.
Hexaploid wheat. Hanxuan 10/Lumai 14 120 DHs evaluated in the field at 2 N levels (2 years) and in hydroponics. An et al. (2006).	Field: NUP. Hydroponics: seedling traits (SDW, RDW, TN), NUP.	1 QTL cluster for NUP, SDW and RDW on 1B.	Field: 9 and 8 QTLs for NUP under low and high N. Hydroponics: 4–5 QTLs for SDW, RDW, TN and NUP. R^2 from 4 to 20%. Co-location of QTLs for NUP in the field and seedling QTLs in hydroponics (1B centromeric, 2DL proximal, 6AL centromeric).
Hexaploid wheat. Arche/Recital: 222 DHs evaluated in the field at 2 N levels (7 environments). Laperche et al. (2007).	GY, yield components, BY, GPC, N harvest index.	1 QTL cluster on 5A near awn inhibitor <i>B1</i> gene.	A total of 37 different chr. regions. Some QTLs stable across N levels and others specific. Semi-dwarf and photoperiod sensitivity genes <i>Rht-B1</i> and <i>Ppd-D1</i> affected all the yield and N-related traits.
Hexaploid wheat. Chinese Spring/SQ1 96 DHs evaluated in pots under greenhouse conditions (near optimal N availability). Habash et al. (2007).	21 growth-related traits, GY, leaf N, GS activity, GPC.	QTLs ($R^2 > 15\%$) for most traits.	Positive association between QTLs for GS-activity and GPC and no association with GY. 2 QTL clusters for GS-activity co-localised with <i>GS2</i> (chr. 2A) and <i>GSr</i> (chr. 4A) genes. Co-localisation of QTLs (6BL, <i>m87p78.5a</i>) for N in peduncle, flag leaf and grain.

Table 18.1A (continued)

Genetic materials and references	Traits	Chromosomes with major QTLs	Other QTLs and comments
Hexaploid wheat. Lovrin10/Chinese Spring 92 DHs at seedling stage in pots (2 P levels). Su et al. (2006).	Soil P deficiency tolerance SDW, TN, SPU, SPUE.	1 QTL for SPU (5D) and 1 for SPUE (5A);	A total of 39 QTLs. 3 QTL clusters on 4B (<i>Xgwm251</i>), 5A (<i>Xgwm271.2</i>), 5D (<i>Xgwm121</i>) with the favourable allele from Lovrin10. QTL clusters on 5A and 5D coincident with the location of <i>VRN</i> -genes.

Abbreviations. ABA: abscisic acid; BY: biomass yield; CFKP: chlorophyll fluorescence kinetics parameters; ChlC: chlorophyll content; CID: carbon isotope discrimination; eOR: cold response; DLT: detached leaf test; DM: dry matter; DSI: drought-susceptibility index; DW: dry weight; FIGD: frost-induced grain damage; FIS: frost-induced sterility; FLS: flag leaf senescence (stay green); FT: frost tolerance; GD: grain dormancy; GF: grain filling; GPC: grain protein content; GrH: growth habit; GS: glutamine synthase; GW: grain weight; GY: grain yield; HD: heading date; KW: mean kernel weight; LTT: low temperature tolerance; NUP: nitrogen uptake; OA: osmotic adjustment; OP: osmotic potential; OPFT: osmotic potential at full turgor; PedL: peduncle length; PEG: poly ethylene glycol; PH: plant height; PHS: preharvest sprouting; RDW: root dry weight; RL: root length; Rno: root number; RPA: root penetration ability; RWC: relative water content; SDW: shoot dry weight; SL: spike length; SPUE: shoot P utilization efficiency; SPU: shoot P uptake; TFID: total frost induced damage (FIS + FIGD); TI: tolerance index; TN: tiller number; VR: vernalization requirement; WS: water stress; WSC: water soluble carbohydrate; WSu: winter survival; WW: well watered.

Table 18.1B Main features of a number of studies targeting the identification of genes for tolerance to abiotic stresses in the Triticeae

Stress and gene	Species	Locus, gene product and gene function
Drought		
<i>HsDREB1A</i> James et al. (2008).	<i>Hordeum spontaneum</i>	<i>CBF/DREB1A</i> gene isolated from xeric <i>Hordeum spontaneum</i> L. Stable <i>HsDREB1A</i> transformants transactivate reporter genes under the control of the stress – inducible <i>HV1</i> and <i>Dhn8</i> promoters.
<i>HV1</i> and <i>HV42</i> Hong et al. (1992); Shen et al. (1993, 2004); Bahieldin et al. (2005).	<i>Hordeum spontaneum</i>	The ABA-response promoter complexes (ABRC) contain a ACGT core element and a coupling element (CE). Both elements are necessary for ABA induction of gene expression. Late Embryogenesis Abundant protein (LEA). Bread wheat transgenics for barley <i>HV1</i> in the field (6 years): improvement of biomass, plant height and grain yield.
<i>wcs120</i> Houde et al. (1992).	Hexaploid wheat	Dehydrin gene isolated from cold-acclimated wheat. Expression specific for cold treatments. Protein accumulation is induced in less than 24 h treatment. The hydrophilic protein is boiling stable.
<i>TaDREB1</i> Shen et al. (2003).	Hexaploid wheat	Isolated from a wheat drought-induced cDNA library. DRE-binding transcription factor. Contains the EREBP/AP2 domain. Induced by drought, low temperature, salinity. The expression of <i>Wcs120</i> and <i>TaDRE1</i> are related.
NF-Y Transcription factor family Stephenson et al. (2007).	Hexaploid wheat	Nuclear Factor Y (NF-Y) is a trimeric complex that binds to the CCAAT box, a ubiquitous eukaryotic promoter element. 37 NF-Y and Dr1 genes identified in wheat. 9 NF-Y and 2 Dr1 genes expressed in wheat leaves are drought responsive.
Extreme temperatures		
<i>Dhn</i> gene family Campbell and Close (1997); Choi et al. (1999, 2000).	Barley	12 components of the <i>Dhn</i> gene family were mapped in Dicktoo x Morex: <i>Dhn 3</i> , <i>Dhn 4</i> , <i>Dhn 5</i> , <i>Dhn 7</i> and <i>Dhn12</i> (6H), <i>Dhn 10</i> (3H), <i>Dhn1</i> , <i>Dhn2</i> and <i>Dhn9</i> (5H). Two <i>Dhn</i> genes (5HL) co-mapped with the winter hardness QTL. Specific PCRs were developed.
<i>Dhn</i> gene family Zhu et al. (2000); Rodriguez et al. (2005).	Barley	Pattern of expression of <i>Dhn</i> genes in plants grown in different climatic conditions. Most of <i>Dhn</i> genes induced by dehydration and freeze-thaw in the field. HMW- <i>Dhn5</i> and LMW- <i>Dhn8</i> induced by low-temperatures only. <i>Dhn13</i> , a new member of the families, is highly expressed with freezing conditions.

Table 18.1B (continued)

Stress and gene	Species	Locus, gene product and gene function
<i>COR14b</i> Crosatti et al. (2003, 2008).	Barley	The cold – regulated protein COR14b encodes a chloroplast-localised protein. The expression of COR14b is controlled by the major QTL region located on chr. 5HL. COR14b correlates with frost resistance only at the earlier field sampling stages and at relatively mild level of controlled hardening.
<i>CBF</i> transcription factor gene family Xue (2002, 2003); Choi et al. (2002); Skinner et al. (2005).	Barley	At least 20 copies of HvCBFs in the barley genome; three subgroups: HvCBF1, HvCBF3, and HvCBF4. <i>HvCBF1</i> : cold-inducible TF containing an AP2 DNA-binding domain. <i>HvCBF2</i> : the CRT/DRE binding capacity of <i>HvCBF2</i> increases at decreasing temperatures. <i>HvCBF3</i> : <i>CBF/DREB1</i> barley homologs were isolated using Arabidopsis <i>CBF/DREB1</i> genes and <i>HvCBF3</i> was mapped in Dicktoo/Morex. HvCBF specifically bound core gene CRT elements.
<i>CBF</i> transcription factor gene family and their regulators Skinner et al. (2006).	Barley	Chr. location of 17 out of 20 <i>CBF</i> genes and of the barley homologs of <i>ICE1</i> and <i>ZAT12</i> genes. 12 <i>CBFs</i> were located on chr. 5HL, 11 of them formed 2 tandem clusters. Based on the map positions, the tandem clusters are candidates only for <i>Fr-H2</i> and not for <i>Fr-H1</i> .
<i>HvCBF</i> , <i>HvMYB</i> Tondelli et al. (2006).	Barley	Cold-response candidate genes isolated and mapped in 3 mapping populations: a cluster of 6 <i>HvCBF</i> genes co-mapped with <i>Fr-H2</i> cold tolerance (5HL). The barley homologs of two Arabidopsis <i>CBF</i> regulators (<i>AtFRY1</i> and <i>AtICE1</i>) mapped outside of the major QTL region. A consensus map of cold and drought QTLs is reported for barley.
<i>HvCBF4</i> Oh et al. (2007).	Barley	Induced by low temperatures. Over-expression in rice increases tolerance to drought, salinity and low-temp. <i>HvCBF</i> activated 15 rice genes (microarray experiment). Some genes are specifically activated by <i>HvCBF4</i> .
<i>Hv-WRKY38</i> Marè et al. (2004).	Barley	TFs early and transiently expressed during exposure to low non freezing temperatures. The gene is ABA independent. Mapped to centromere of chr. 6H.
<i>Ltp-6</i> Federico et al. (2005).	Barley	<i>Ltp6</i> encodes a lipid transfer protein. The gene is responsive to cold, salt and ABA treatments. The promoter is specific of the epicarp of the seed.
<i>HvMC1</i> Philipps et al. (2006).	Barley	Cold and light stress mediated activation of the gene belonging to the mitochondrial carrier proteins. The regulatory path where the gene is included could involve a chloroplast derived signal.

Table 18.1B (continued)

Stress and gene	Species	Locus, gene product and gene function
<i>Fr1</i>	Hexaploid wheat	The <i>Vrn1</i> and <i>Fr1</i> loci are closely linked on the distal portion of the long arm of 5AL near to RFLP markers <i>Xpsr426</i> , <i>Xcdo504</i> and <i>Xwg644</i> .
Galiba et al. (1995). <i>wcs19</i> and <i>wcor15</i>	Hexaploid wheat	Members of the wheat cold-responsive gene family induced by low-temperatures. 5 copies of <i>wcor15</i> on chr. group 2. <i>wcor15</i> is ABA-independent. The promoter contains the CRT/DRE motifs and induces expression at low temperature.
Chauvin et al. (1993); Takumi et al. (2003); Shimamura et al. (2006).		
<i>Vrn1</i>	Hexaploid wheat	<i>Vrn1</i> is the MADS-box gene <i>AP1</i> , regulated by vernalization treatments in apices and leaves. Mutation in the promoter region is the determinants of the spring/winter habit. Strong epistatic interaction between <i>Vrn1</i> and <i>Vrn2</i> . Homologs: <i>Vrn-H1</i> (Sgh2) in barley (5H) and <i>Sp1</i> in rye (5R).
Galiba et al. (1995); Laurie et al. (1995); Plaschke et al. (1993); Yan et al. (2003).	Hexaploid wheat	Cold-responsive gene. Signal peptide.
<i>wcor14</i>	Hexaploid wheat	
Tsvetanov et al. (2000).		
Other wheat <i>Cor/Lea</i> genes (<i>Wrab17</i> , <i>Wrab18</i> , <i>Wrab19</i> , <i>Wcor825</i>)	Hexaploid wheat	Specifically induced by low temperatures.
Tsuda et al. (2000); Kobayashi et al. (2004).		
<i>wcbf/dreb2</i> (homolog of CBF/DREB1)	Hexaploid wheat	The wheat <i>wcbf2</i> gene activates the <i>Cor/Lea</i> genes. Transformation of <i>wcbf2</i> in tobacco alters the expression pattern of cold-responsive genes and increase freezing tolerance.
Kume et al. (2005); Kobayashi et al. (2005); Takumi et al. (2008).		
15 <i>CBF</i> genes	Hexaploid wheat	Up to 25 <i>CBF</i> genes, classified into 10 groups, present in wheat species. 5 groups are low temperature inducible in winter wheat. All <i>TaCBF</i> genes display a transitory induction profile.
Badawi et al. (2007).		
<i>wlt10</i> (LT-COR protein family)	Hexaploid wheat	Specifically induced by low temperatures. Quantitative expression correlates with freezing-tolerance. Peak of mRNA accumulation when the maximum level of freezing tolerance is attained. Signal peptide.
Ohno et al. (2001).		

Table 18.1B (continued)

Stress and gene	Species	Locus, gene product and gene function
<i>Wpi-6</i> Imai et al. (2005).	Hexaploid wheat	Plasma membrane protein with membrane spanning domain. Strong inductions observed at the early stages of cold acclimation, as well as salinity and ABA. It may have a protective role.
<i>Waox1a</i> Sugie et al. (2006).	Hexaploid wheat	<i>Waox1a</i> (mitochondrial alternative oxidase) overexpressed in Arabidopsis enhances recovery of the total respiration activity under low temperature. Levels of ROS decreased in the transgenics under low temperature stress.
<i>TaCDPK1</i> Martinez-Noel et al. (2007).	Hexaploid wheat	<i>TaCDPK1</i> (Ca ⁺⁺ -dependent protein kinase) is upregulated by sucrose treatment of excised leaves.
<i>CDPK</i> Li et al. (2008).	Hexaploid wheat	20 CDPK (calcium-dependent protein kinases) genes respond to cold, H ₂ O ₂ , salt, drought, ABA, GA, powdery mildew. Each CDPK gene responds to multiple treatments.
<i>Ta-A-Inv</i> Vargas et al. (2007).	Hexaploid wheat	Alkaline invertase expression up-regulated in wheat leaves after low-temperature/osmotic treatment. The enzyme hydrolyzes sucrose and contribute to a more efficient cytosolic sucrose hydrolysis during environmental stresses.
alpha-tubulin gene family Farajalla and Gulick (2007).	Hexaploid wheat	15 genes in 5 groups, each with the presence of the 3 homologues. Cold acclimation induces an initial decrease in expression level, followed by an increase with different patterns of reinduction.
<i>HSP101</i> gene family Gulli et al. (2007).	Durum wheat	Heat Shock Protein HSP101 has a central role in heat stress survival. Four <i>HSP101</i> members were isolated and characterized: different regulation and distinct roles in stress response and thermotolerance acquisition.
<i>Cbf1A</i> , <i>Cbf1C</i> , <i>Cbf7</i> Miller et al. (2006).	<i>Triticum monococcum</i>	<i>CBF</i> genes mapped at the 5A ^m chromosome regions underlying <i>Ftr-A^m2</i> .
<i>LHC II</i> , <i>rbcs</i> and <i>rbcl</i> , <i>pmr</i> . <i>Scpmr</i> In et al. (2005).	Rye	Genes strongly expressed in 4°C-cold hardened leaves; reduced expression during dehardening. <i>Scpmr</i> does not respond to heat or cold exposure in darkness. Antioxidative enzyme peptide methionine sulfoxide reductase.
<i>Lls1</i> In et al. (2005).	Rye	The lethal spot 1 gene strongly induced during the dehardening (22°C for 3 days). Plasma-membrane localized H ⁺ -ATPase, protein disulfide isomerase.

Table 18.1B (continued)

Stress and gene	Species	Locus, gene product and gene function
Salinity		
<i>HVP1</i> , <i>HVP10</i> , <i>HvVHA-A</i> , <i>HvNHX1</i>	Barley	Transcript levels of <i>HVP1</i> , <i>HVP10</i> , <i>HvVHA-A</i> , <i>HvNHX1</i> increase with salt stress. <i>HVP1</i> and <i>HvNHX1</i> also respond to osmotic stress. Expressions of these genes are different as to root and shoot organs. <i>HvNHX1</i> is mainly expressed in roots and not in shoot. ABA and auxin markedly induce the expression of <i>HVP1</i> and <i>HvNHX1</i> , but not of <i>HVP10</i> .
<i>HvHKT1</i> and <i>TaHKT1</i>	Barley	HKT1 is a high-affinity root K^+ transporter that functions as high-affinity Na^+ uptake system of roots of wheat and barley. High affinity Na^+ uptake appears soon after K^+ depletion.
<i>HvA1</i>	Barley	The barley <i>HvA1</i> conferred greater salt tolerance to adult plants in transformed transgenic oat (<i>Avena sativa</i> L.).
Oraby et al. (2005).		A member of the plasma membrane and vacuolar membrane Na^+/H^+ antiporter proteins. <i>HvNHX2</i> is expressed in roots, stems and leaves of barley seedlings. The amount of <i>HvNHX2</i> protein is not correlated to the Na^+/H^+ exchange rate at different salt stresses.
<i>HvNHX2</i>	Barley	
Vasekina et al. (2005); Ershov et al. (2007).		
<i>HvPIP</i> genes	Barley	<i>HvPIP</i> (aquaporins) transcripts slightly increase under salt stress.
Katsuhara and Shibasaki (2007).		
<i>HvProT</i>	Barley	The proline transporter is highly expressed in the apical region of barley roots under salt stress. Free proline transport is a component of cell wall synthesis in the apical region of barley roots.
Ueda et al. (2007).		
<i>HvRAF</i>	Barley	<i>H. vulgare</i> root abundant factor is an ethylene response- type factor more abundant in roots than shoots. It contains the <i>APETALA2/ERF</i> DNA-binding domain. Overexpression of <i>HvRAF</i> in Arabidopsis enhances pathogen resistance and salt tolerance.
Jung et al. (2007).		
Root Plasma Membrane Transporters	Barley	Root plasma membrane (PM) transporters control K^+/Na^+ homeostasis in salt-stressed barley. Salt tolerant barleys have superior ability to pump Na^+ from cytosol to external medium. Candidate gene for active Na^+ extrusion: PM-bound Na^+/H^+ antiporter, SOS1. Tolerance correlates with higher H^+ -ATPase activity in roots.
Chen et al. (2005, 2007a,b).		
<i>HsDREB1A</i>	Barley	Transgenic <i>Paspalum notatum</i> Flügge with stress-inducible expression of HsDREB1A are resistant to severe salt stress and repeated cycles of severe dehydration stress.
James et al. (2008).		

Table 18.1B (continued)

Stress and gene	Species	Locus, gene product and gene function
<i>TmHKT7-A1</i> , <i>TmHKT7-A2</i> , <i>Nax1</i> Huang et al. (2006).	Hexaploid wheat, <i>T. monoccum</i>	2 putative sodium transporter genes (<i>TmHKT7</i>), located on the same BAC contig at 145 kb from each other, are candidate cosegregating with <i>Nax1</i> . <i>TmHKT7-A2</i> is expressed in roots and leaf sheaths of salt tolerant Line 149. <i>TmHKT7-A2</i> could control Na ⁺ unloading from xylem in roots and sheaths.
<i>TaHKT1;5-D</i> , <i>Kna1</i> , <i>Nax2</i> James et al. (2006); Byrt et al. (2007)	Durum wheat	High affinity K ⁺ transporter involved in Na-tolerance. Facilitate Na ⁺ unloading from xylem, leading to reduced Na ⁺ accumulation in the shoots. No allelic variation known. The <i>Nax2</i> gene has a similar function to bread wheat <i>Kna1</i> . <i>Nax2</i> and <i>Kna1</i> are associated with HKT1;5 gene.
<i>TNHX1</i> and <i>TVPI</i> Brimi et al. (2007).	Hexaploid wheat	Overexpression of wheat vacuolar Na ⁺ /H ⁺ antiporter <i>TNHX1</i> and H ⁺ -pyrophosphatase (H ⁺ -PPase) <i>TVPI</i> improve salt-and drought- stress tolerance in <i>Arabidopsis thaliana</i> plants. Transgenics grow well at 200 mM NaCl and under water-deprivation. The osmotic adjustment is enhanced. The vacuolar solute accumulation increase.
<i>TaNHX2</i> Yu et al. (2007).	Hexaploid wheat	cDNA isolated from seedlings treated with NaCl. Salt induces a steady accumulation in roots and leaves. <i>TaNHX2</i> also induced by ABA, PEG and cold. Expression of <i>TaNHX2</i> in yeast improves the tolerance of the <i>Anhx1</i> .
<i>HKT1</i> , <i>SOS1</i> , <i>NHX1</i> , <i>NHX5</i> , <i>AVP1</i> , <i>AVP2</i> Mullan et al. (2007).	Hexaploid wheat, <i>L. elongatum</i>	Wheat- <i>L. elongatum</i> amphiploid and wheat- <i>L. elongatum</i> disomic substitution lines. Arabidopsis-rice-wheat-wheatgrass ortholog genes were isolated and mapped to wheat chromosomes. <i>SOS1</i> gene was mapped on the homoeologous group 3, short arm, where the <i>L. elongatum</i> Na ⁺ exclusion factor resides.
Ferredoxin-thioredoxin reductase In et al. (2005).	Rye	Genes strongly expressed in 4°C-cold hardened leaves followed by a maintenance of the expression level during the dehardening (22°C for 3 days).
Toxic ions <i>HvMATE</i> / <i>HvAACT1</i> (<i>Alp</i>) Wang et al. (2007); Furukawa et al. (2007).	Barley	Al-toxicity tolerance. Citrate efflux transporter (Aluminium activated citrate transporter); citrate is pumped out of the root and binds/detoxifies Al ³⁺ . MATE proteins are known to facilitate citrate efflux from Arabidopsis and other species.

Table 18.1B (continued)

Stress and gene	Species	Locus, gene product and gene function
<i>TaALMT1</i> (<i>AltBH</i>) Raman et al. (2005); Sasaki et al. (2006).	Hexaploid wheat	Al-toxicity tolerance. Malate efflux transporter (Aluminium activated malate transporter); malate is pumped out of root and binds/detoxifies Al^{3+} . Expression is positively associated with the copy number of a promoter repeat element. Sequence variation in the 1,000 bp upstream correlates with <i>ALMT1</i> expression level and Al resistance.
<i>ScALMT1</i> (<i>Alt4</i>) Fontecha et al. (2007).	Rye	Al-toxicity tolerance. Malate efflux transporter (putative Aluminium activated). Al-inducible mRNA levels are higher in tolerant genotypes (5-fold). <i>ScALMT1</i> gene is expressed in the root.
<i>BOR1</i> ortholog; efflux transporter (<i>Bot1</i>) Sutton et al. (2007).	Barley	Boron tolerance. <i>Bot1</i> mapped using a candidate gene approach from the <i>Arabidopsis BOR1</i> efflux transporter ortholog. The rice syntenic region lacked a <i>BOR1</i> ortholog. The transporter is mainly expressed in young cells of the roots and could serve the dual role of limiting symplastic boron concentration and helping to maintain high Boron supply. The tolerant parent Sahara contains four times as many <i>Bot1</i> gene copies.
10 NRAMP-transporter genes Finkemeier et al. (2003).	Barley	The inter-relation between N availability and cadmium toxicity was studied in roots by comparing the transcription profiles of 10 genes involved in antioxidant defence and cadmium transporter (NRAMP-transporter) genes at different N levels. NRAMP transporter was down-regulated in the presence of Cd at adequate N levels.
Low nutrients – Nitrogen and phosphorus		
<i>HvNRT2.1</i> and <i>2.2</i> ; <i>HvNRT2.3</i> and <i>2.4</i> Trueman et al. (1996); Vidmar et al. (2000); Tong et al. (2005). <i>HvPhl1.1</i> and <i>HvPhl1.2</i> Schunmann et al. (2004a,b).	Barley	Inducible high-affinity NO_3^- transporters. The 4 high-affinity nitrate transporters have different transcription patterns in response to variation in nitrate and ammonium supply. <i>HvNRT2</i> transcripts were undetectable in leaf tissues. Barley phosphate transporter genes (gene family, 8 members) have a relevant role in the uptake of phosphate by root hairs from soil solution. Within the promoter sequence, a single PIBS element increases the response to P deprivation.

Table 18.1B (continued)

Stress and gene	Species	Locus, gene product and gene function
<i>GSI</i> , <i>GS2</i> , <i>GDH</i> and <i>NADH</i> and <i>NADPH</i> nitrate reductase	Cereals, including hexaploid wheat	10 <i>GS</i> genes exist in wheat (<i>GSIa</i> , <i>b</i> , <i>c</i> , <i>GS2a</i> , <i>b</i> , <i>c</i> , <i>GSI1</i> , <i>2</i> , <i>GSI1</i> , <i>2</i>). Feedback regulation of aminoacids and nitrate reductase activity. Cytosolic isozymes <i>GSI</i> and <i>GSIr</i> have a major role in assimilating ammonia during N remobilization to the grains. Activity of key-N metabolic enzymes (<i>GSI</i> , <i>GS2</i> , glutamate dehydrogenase) could be manipulated to increase NUE.
Miflin and Habash (2002); Fan et al. (2006); Bernard et al. (2008).		
NAC transcription factor; NAM-B1 (<i>Gpc-B1</i>)	Durum wheat	The functional wild wheat (<i>T. dicoccoides</i> Korn.) allele accelerates senescence and increases nutrient (including nitrogen, zinc, iron) remobilization from leaves to developing grains. Modern wheat carries a non-functional NAM-B1 allele.
Uauy et al. (2006). <i>WPBF</i>	Hexaploid wheat	Wheat DOF transcription factor activates storage protein genes during seed development. <i>WPBF</i> is expressed in several tissues and TaQM protein, isolated from roots, interacts with the DOF domain. The promoter activity of <i>WPBF</i> is observed in seeds and vascular system of transgenic Arabidopsis.
Dong et al. (2007).		
<i>TaPT2</i>	Hexaploid wheat	Expression pattern of a putative wheat high-affinity phosphate transporter in roots and leaves under P-deficiency. Identification of a promoter fragment responding to P-deficient conditions in Arabidopsis. Identification of conserved regions and candidate regulatory motifs with transporter promoters from barley, rice and Arabidopsis.
Tittarelli et al. (2007).		

Pre-harvest sprouting

Vp-1B (*Viviparous-1B*)
Wilkinson et al. (2002, 2005);
Yang Y. et al. (2007a,b);
Xia et al. (2008).

AB3/Vp-1 gene family has been suggested as a potential candidate for PHS and grain dormancy. Wheat *Vp-1* homologues (*Vp-1A*, *Vp-1B*, *Vp-1D*) have been mapped to the long arms of group 3 chromosomes (conserved synteny with *Vp1* in maize and *Osvp1* in rice). The reduced activity of the protein in maturing grains may contribute to PHS. At *Vp-1B* (3BL), several *Vp-1* alleles are present in the cultivated hexaploid wheat germplasm. The phenotypic effect of each allele has been evaluated with an association mapping approach.

Table 18.2 Main features of microarray studies in the Triticeae targeting the expression of genes influenced by abiotic stresses

Species and stress	Experimental conditions and		Main results
	Objective cross, sample and platform	sampling protocol	
Barley Drought and salt stress. Ozturk et al. (2002).	Monitoring responses to drought and salinity in barley.	Drought stress: <ul style="list-style-type: none"> • pots with sand • 3 weeks after germination • plants removed from sand and left on paper. Tissue collected at 6 and 10 h of dehydration. 	Transcripts with significant up-regulation under dehydration: jasmonate-responsive, metallothionein-like, late-embryogenesis-abundant and ABA-responsive proteins. Up-regulation (drought and salt stresses) was restricted to ESTs for metallothionein-like and LEA proteins; increases in ubiquitin-related transcripts (salt stress). Ninety-nine genes up- or down-regulated in at least one condition.
	<ul style="list-style-type: none"> • cv. Tokak • leaf and root • cDNA microarray: 1,463 clones from barley related to drought-stress. 	Salt stress: <ul style="list-style-type: none"> • hydroponic tanks • 3 weeks after germination • 150 mM NaCl. Tissue collected at 24 h. 	
Barley Drought, cold, salinity, high light, Cu, heat + drought, cold + Cu. Atienza et al. (2004).	Identify genes involved in barley adaptation to abiotic stress.	<ul style="list-style-type: none"> • pots covered by Whatman paper • 7-d old seedling • cold: 4°C for 53 h (light or darkness). Drought: new pods with dry whatman paper at 18°C for 5 and 10 h or 30–38°C for 5 h. 	Two abiotic factors simultaneously evidenced different responses compared to the corresponding single stress conditions.
	<ul style="list-style-type: none"> • cv. Nure • leaf • cDNA microarray: 300 clones from barley and wheat related to abiotic stress. 	Salinity: 175 mM NaCl for 24 h. High light: 1.15 mmol m ⁻² s ⁻¹ for 10 h. Cu: 1 mM CuCl ₂ for 24 h.	
Barley Drought Talamè et al. (2007).	Compare transcript profiles induced by slow and rapid drought-stresses.	<ul style="list-style-type: none"> • pots with compost, vermiculite and sand • four-leaf stage plants 	Changes in expression profiles according to the dehydration treatment. A 10% of transcripts had similar changes at the two water stresses; a number of transcripts shared a common trend in both experimental conditions.
	<ul style="list-style-type: none"> • cv. Er/Apm • leaf • cDNA microarray (Ozturk et al. 2002). 	<ul style="list-style-type: none"> • dehydration shock treatment: dried with paper towels and left at 27°C for 6 h. Drought stress treatment: 7 and 11 days withholding water. 	

Table 18.2 (continued)

Species and stress	Objective cross, sample and platform	Experimental conditions and sampling protocol	Main results
Wheat Drought Way et al. (2005).	Identify genes involved in wheat adaptation to water deficit. • cv. Kennedy • leaf • cDNA microarray: 300 clones from wheat related to drought-stress (subtractive libraries).	• pots containing soil • water withholding for 12 days starting at the boot stage (DC45) until approximately anthesis • tissue collected every day from the beginning of the imposition of the stress.	Many genes up-regulated in this long-term stress array were in common with those already identified in short term stress. Decreases were detected for the rate of photosynthesis and starch accumulation in leaves.
Wheat Drought Xue et al. (2006).	Identify genes differentially regulated in wheat lines during drought stress. • cv. Quarrion and Genaro81 (10 BC ₂ F _{5:7}) • leaf • cDNA microarray: 16,000 unique wheat ESTs related to abiotic stress.	• (i) Field trial and (ii) controlled environment (pots containing Diatomite) • (i) prior to the onset of drought and (ii) booting stage • (ii) withholding water for 10–12 days.	Ninety-three genes differentially expressed between high and low transpiration efficiency progeny lines were identified. One fifth of these genes were markedly responsive to drought stress.
Wheat Drought Mohammadi et al. (2007).	Monitoring responses to drought in dehydration-tolerant wheat. • cv. Opata and [Altair 84/A. squarrosa (TAUS)//Opata] • root • oligo array: 19,000 oligonucleotide probes related to abiotic stress.	• pots containing turf surface MPV medium • water withholding beginning on 21 day after seeding • Tissue collected after 36 h of the water stress treatment.	The major classes of dehydration-responsive genes were generally similar to those observed previously in other species and osmotic stresses. A novel, dehydration-induced putative AP2/ERF transcription factor, and a dehydration induced “little protein” were identified.

Table 18.2 (continued)

Species and stress	Objective cross, sample and platform	Experimental conditions and sampling protocol	Main results
Barley Salt and osmotic stress Ueda et al. (2004).	Dissect genes responding to osmotic potential changes under salt stress. • cv. Haruna-niyo • leaf and root • cDNA microarray: 460 barley salt-responsive genes (differential display).	• hydroponic trays • three-leaf stage plants • PEG (equivalent to 200 mM NaCl). Tissue collected 1 and 24 h after the osmotic stress treatment.	Out of 52 genes differentially expressed under osmotic stress, 11 were regulated in a virtually identical manner under salt stress; this suggests that several cellular responses are mediated by osmotic potential changes.
Barley Salt stress Walia et al. (2005).	Understand the early responses of genes to salinity stress at seedling stage. • cv. Morex • shoot • oligo array: 22-k Barley1 Affymetrix.	• hydroponic trays • day 16 after germination • 100 mM NaCl by increments of 25 mM NaCl per day.	Transcriptome analysis indicates induction of genes involved in jasmonic acid biosynthesis and genes known to respond to jasmonic acid treatment.
Barley Salt stress Ueda et al. (2006).	Understand the molecular bases that differs barley and rice ability to tolerate salt stress. • cv. Haruna-niyo and rice cv. IR64 • leaf and root • cDNA microarray: 460 salt-responsive genes of barley (differential display).	• hydroponic trays • two-leaf stage plants • 200 mM NaCl for barley and 150 mM NaCl for rice. Tissue collected after 1 and 24 h of salt stress treatment.	Transcripts for plasma membrane protein 3 and inorganic pyrophosphatase were up-regulated in both species, but only transiently in rice. Adaptive mechanisms for regulating ion homeostasis are partly conserved in the two species, but it seems that rice cannot sustain cellular ion homeostasis for a long time like barley.

Table 18.2 (continued)

Species and stress	Objective cross, sample and platform	Experimental conditions and sampling protocol	Main results
Barley Salt stress Walia et al. (2007a).	Elucidate the genetic basis of salt tolerance. ● cv. Golden Promise and Maythorpe ● leaf and root ● oligo array: 22-k Barley1 Affymetrix.	● hydroponic trays ● four-leaf stage plants ● 150 mM NaCl. Tissue collected at day 9 of the salt stress treatment.	The response of the two genotypes to salinity stress is quite different, particularly the jasmonic acid pathway regulation. No known gene with a functional association with the favourable Na ⁺ homeostasis maintained by Golden Promise were identified.
Barley Salt stress Walia et al. (2007b).	Compare expression after salinity stress, Jasmonic Acid (JA) treatment and JA pre-treatment followed by salinity stress. ● cv. Golden Promise ● growing point tissue ● oligo array: 22-k Barley1 Affymetrix.	● hydroponic trays ● day 8 (JA) and 10 (salt) after germination ● final concentration of 12 mM of JA; a salinity stress of ~18 dS m ⁻¹ . Tissue collected on day 14 after germination.	A considerable overlap between genes regulated by salinity stress and JA application was detected. Three JA-regulated genes, arginine decarboxylase, rubisco activase and apoplactic invertase are possibly involved in salinity tolerance mediated by JA.
Wheat Salt stress Kawaura et al. (2006).	Examine the transcriptional response of wheat to salt stress. ● cv. Chinese Spring ● shoot and root ● oligo array: 22-k oligoarray (probed selected based on annotation).	● vermiculite ● fourteen-day-old ● 150 mM NaCl. Tissue collected after 0, 6 and 24 h of the salt stress treatment.	Identified 1,811 genes whose expressions changed more than 2-fold in response to salt. These included genes known to mediate response to salt, as well as unknown genes.

Table 18.2 (continued)

Species and stress	Objective cross, sample and platform	Experimental conditions and sampling protocol	Main results
Wheat			
Salt stress			
Mott and Wang (2007).	Identify differentially expressed genes in roots and leaves under salt stress. <ul style="list-style-type: none"> cv. Chinese Spring, W4909, W4910, AJDAj5 and Ph¹ shoot and root oligo array: GeneChip1 Affymetrix. Monitor transcriptional alterations during high-temperature injury. <ul style="list-style-type: none"> cv. Haruna-nijyo panicle and shoot oligo array: 22-k Barley1 Affymetrix. 	<ul style="list-style-type: none"> silica sand Three days after planting Progressively increasing salt concentration (initial electrical conductivity of 6 dS/m and final of 30 dS/m). Tissue collected at day 39 of the salt stress treatment. growth cabinet Panicle: <ul style="list-style-type: none"> Four-leaf stage plants 30°C (day), 25°C (night). Panicle collected at 3, 5, 10, and 20 mm in length Shoot: <ul style="list-style-type: none"> 7 days after germination 30°C (day), 25°C (night). Tissue collected 5 days after stress treatment. <ul style="list-style-type: none"> soil/sand 7 days after planting cold acclimatized in a daily regime of 8 h light at 3°C and 16 h dark at 2°C. Tissue collected at day 4 after cold stress treatment. 	<p>Salt treatment responsive gene expression profiles common to all lines provide a list of candidate salt-tolerance genes in wheat. A gene for tonoplast aquaporin and a gene for putative K channel protein were identified as candidate genes for salt tolerance. Abiotic or biotic stress related genes were equally or more dominantly upregulated in the seedlings exposed to high temperatures compared with the panicles. In contrast, certain genes associated with histones, DNA replication initiation, mitochondria, and ribosomes were specifically repressed in the exposed panicles.</p> <p>Numerous previously reported cold-related genes were differentially expressed in treated and untreated plant. Many of the genes found do not have an assigned function.</p>
Barley			
Cold stress			
Faccioli et al. (2002).	Identify genes for cold stress response. <ul style="list-style-type: none"> Nure, Tremois, Proctor, Nudinka and <i>Hordeum spontaneum</i> HS65 leaf cDNA microarray: 900 barley clones related to drought stress. 		

Table 18.2 (continued)

Species and stress	Objective cross, sample and platform	Experimental conditions and sampling protocol	Main results
Wheat	Monitor gene expression during cold acclimation.	• vermiculite	Transcript levels of more than 300 genes were altered by cold. These include genes that encode potential regulatory proteins, proteins that act in plant metabolism and proteins involved in photosynthesis.
Cold stress	• cvs. Glenlea and Norstar	• 7 days after planting	
Gulick et al. (2005).	• shoot	• seedlings were cold acclimated at 6±1°C during the day and 2±1°C during the night. Tissue collected at day 1, 6 and 36 after cold stress treatment.	
	• cDNA microarray: 1,184 clones representing 947 genes related to cold stress.	• soil/vermiculite	
Wheat	Study signaling and regulatory pathways of the changes in gene expression associated with acquired freezing tolerance.	• 7 days after planting	Detected over 450 genes regulated by cold treatment and differentially regulated between spring and winter cultivars. Of these, 130 are signaling or regulatory gene candidates, including transcription factors, protein kinases, ubiquitin ligases and GTP, RNA and calcium binding proteins.
Cold stress	• cvs. CDC Clair and Quantum	• seedlings were cold acclimated by lowering the growth temperature to 4°C. Tissue collected 6 h, 1, 3, 6, and 14 days after cold stress treatment.	
Monroy et al. (2007).	• shoot		
	• cDNA microarray: 5,740 clones enriched for signal transduction and regulatory genes.		
Barley	Study genes involved in the methionine cycle.	• hydroponic trays	Expression of genes involved in methionine cycle increased in Zn-deficient and Fe-deficient barley roots, thus allowing the plant to meet its demand for methionine, a precursor of murgineic acid (MA). Transcripts involved in MA synthesis detected in Zn-deficient shoots (not in Fe-deficient).
Fe and Zn deficiency	• Barley cv. Ehimehadaka no. 1	• six days after planting	Ca. 200 clones identified as Fe-deficiency-inducible genes.
Suzuki et al. (2006).	• shoot and root	• omitting Fe-EDTA or ZnSO ₄ from the culture medium. Tissue collected at day 14 of the Zn and Fe deficiency treatment.	Polar vesicle transport is involved
	• cDNA microarray: 8,987 rice ESTs.		
Barley	Examine gene expression profiles in barley roots during Fe-deficiency stress.	• hydroponic trays	
Fe deficiency		• three-leaf stage plants	
Negishi et al. (2002).			

Table 18.2 (continued)

Species and stress	Objective cross, sample and platform	Experimental conditions and sampling protocol	Main results
Barley Fe deficiency Ogo et al. (2006).	<ul style="list-style-type: none"> ● cv. Ehimehadaka no. 1 ● root ● cDNA microarray: 8,987 rice ESTs. 	<ul style="list-style-type: none"> ● omitting Fe from the culture medium. Tissue collected at day 14 of the Fe deficiency treatment.. 	in the diurnal secretion of mugineic acid family phytosiderophores (MAs).
	Clarify the molecular mechanism that regulates iron (Fe) acquisition.	<ul style="list-style-type: none"> ● hydroponic trays ● three weeks after germination ● omitting Fe(III)-EDTA from the culture medium. Tissue collected at day 0, 1, 2, 3, 5, and 7 after Fe-deficiency treatment. Fe was re-supplied at day 7, and the plants were harvested at day 11. 	Twenty-one genes for proteins that function in gene regulation were induced by Fe deficiency. Of these genes, a transcription factor gene, named OsIRO2, is probably involved in the regulation of gene expression under Fe-deficient conditions.
	<ul style="list-style-type: none"> ● cv. Ehimehadaka no. 1 and rice cv. Nipponbare ● shoot and root ● oligo array: rice 22 K oligo-DNA microarray (Agilent). 	<ul style="list-style-type: none"> ● sand and perlite/vermiculite ● germinating seeds ● watering with low N solution. Tissue collected at anthesis. 	Fructan accumulation in stems was accompanied by elevated transcripts for a suite of fructosyltransferases (FTs) and for a fructan 6-oxohydrolase (6-FEH) in the low N compared to high N stems. Transcripts for the WPK4 kinase increased in response to sucrose, suggesting a role for this kinase in C storage metabolism in the reproductive wheat stems grown in low N.
Wheat N deficiency Ruuska et al. (2008).	Investigate long-term adaptation to nitrogen (N) limitation of wheat.		
	<ul style="list-style-type: none"> ● cv. Janz ● developing heads, peduncles, flag leaf blades, penultimate internodes, lower stems and lower leaves 		
	<ul style="list-style-type: none"> ● cDNA microarray: 36,000 unique sequences. 		
Barley Al stress Furukawa et al. (2007).	Identify genes that characterize Al-resistant cultivars.	<ul style="list-style-type: none"> ● hydroponic trays ● four-day-old seedlings ● 1.0 mM CaCl₂ solution containing 0 or 5 mM Al for 6 h. 	By using a combination of positional cloning and microarray analysis a gene responsible for Al-induced
	<ul style="list-style-type: none"> ● cv. Murasakimochi and Morex (F₄) ● root apices 		

Table 18.2 (continued)

Species and stress	Objective cross, sample and platform	Experimental conditions and sampling protocol	Main results
Wheat Al stress Xiao et al. (2005).	<ul style="list-style-type: none">● oligo array: 22-k Barley1 Affymetrix. Identify ESTs in roots (aluminum stress).	<ul style="list-style-type: none">● Metro-Mix 360 soil mix● nine days old● 0.3 mmol/L AlK(SO₄)₂ into the culture solution. Tissue collected 48 h after Al treatment.	secretion of citrate has been cloned. 25 ESTs were upregulated in tolerant NIL in response to Al. These ESTs showed homology with genes involved in signal transduction, oxidative stress alleviation, membrane structure, Mg ²⁺ transportation, and other functions.
Wheat Al stress Guo et al. (2007).	<ul style="list-style-type: none">● cv. Century and its NIL OK91G106● root● cDNA microarray: 614 clones from wheat Al-stressed root (subtractive libraries). Understand the mechanisms of aluminum tolerance in wheat.	<ul style="list-style-type: none">● hydroponic trays● Four days old seedling● 0–15 mg/l Al³⁺ by adding AlK(SO₄)₂ into the culture solution. Tissue collected at 0–15 min, 1–3–6 h, 1–3–7 d after Al treatment.	Fifty-seven genes were differentially expressed for at least one time point of Al treatment. Among them, 28 genes were up-regulated in the tolerant line. Genes related to senescence and starvation of nitrogen, iron, and sulfur were highly expressed in the susceptible line.

production of large experimental populations (Buckler and Thornsberry 2002; Ersoz et al. 2007). Important factors to consider for best deploying association mapping are the level of linkage disequilibrium (LD) among the investigated accessions and the presence of population structure that could greatly increase false-discovery rate (Flint-Garcia et al. 2003; Ersoz et al. 2007). Populations characterised by high LD (> 1 cM) are more suitable for a genome-wide search, particularly when the panel of accessions has been profiled at a rather limited number of loci (Maccaferri et al. 2005; Breseghello and Sorrells 2006; Rostoks et al. 2006; Bagge et al. 2007; Crossa et al. 2007; Somers et al. 2007). Conversely, the validation of the role of a candidate sequence requires the utilisation of panels with much lower LD (< 0.02 cM; Salvi et al. 2007), hence a much higher level of genetic resolution, a condition that is more easily attained in allogamous species.

QTL cloning provides an entry point to more effectively exploit sequence variability at target loci and to unlock the allelic diversity present in germplasm collections and wild relatives. Both positional cloning and high-resolution association mapping enable us to identify the most promising candidate sequence for a target QTL (Salvi et al. 2007). Genes influenced by abiotic stress (Tables 18.1B and 18.2) and EST collections from stressed tissues provide a valuable source for the construction of functional maps that can facilitate the identification of candidates for QTLs. In barley, 2,000 genome-wide distributed SNP markers have been developed and mapped from a selection of genes based on their differential transcription response to abiotic stress (Rostoks et al. 2005). Specific efforts toward enriching Triticeae linkage maps with function-specific genes have been undertaken and in some cases have been utilised for QTL analysis (Andersen and Lubberstedt 2003; Diab et al. 2004; Tondelli et al. 2006; Rostoks et al. 2005; Stein et al. 2007). A number of forward- and reverse-genetics approaches allow for a genome-wide functional screen. Among these, RNAi (Waterhouse and Helliwell 2003; Travella et al. 2006; Fu et al. 2007), TILLING and EcoTILLING provide valuable opportunities to elucidate the functional role of target sequences (Comai et al. 2004; Till et al. 2007).

18.2.1 Candidate Gene Approach

The candidate gene approach (CGA) provides a shortcut from the tedious procedures entailed by positional cloning. However, like any other shortcut, the CGA has shortcomings. Because the CGA relies on the information available on open reading frames (ORFs), its effectiveness is reduced when the QTL effect is due to a polymorphism at a *cis*-acting, apparently non-coding sequence that in some cases can be several kb from the effector gene modulating the QTL effect, as recently shown in maize for *Vgt1*, a QTL for flowering time located at ca. 70 kb upstream of the effector gene encoding for an *Apetala 2*-like factor (Salvi et al. 2007). Therefore, as recognised by Rafalski and Morgante (2004),

the identification of regulatory regions often quite distant from the effector genes indicates that the selection of a candidate sequence to be tested for association mapping with a phenotype is by no means a trivial undertaking if the genomic scan aims to be comprehensive. Although the CGA can be deployed also with no prior knowledge about QTLs for the target trait (for an example see Yamasaki et al. 2005), its application usually focuses on genes mapped within the support interval of QTLs for which a plausible cause-effect relationship can be hypothesised between the target trait and the function of the candidate. The identification of suitable candidate genes and the elucidation of their function can be facilitated by combining the “-omics” platforms (e.g. transcriptomics, proteomics, metabolomics, etc.) with laser capture microdissection, a technique that provides unprecedented levels of functional resolution at the tissue level, down to a single-cell layer (Nakazono et al. 2003; Nelson et al. 2006). The technique has already been used in the Triticeae to investigate aspects of plant-pathogen interactions (Wise et al. 2007).

18.2.2 Exploiting the “-omics” Platforms

The functional basis of a number of cloned plant QTLs relates to differences in their expression level (Salvi and Tuberosa 2005). Along this line, an interesting application of transcriptome analysis is the identification of the so-called eQTLs, namely QTLs that control the level of expression (hence the “e”) of a particular gene (Kirst and Yu 2007; Potokina et al. 2008). An eQTLs that co-segregates with the locus for which expression was analysed indicates allelic differences at *cis*-regulatory regions, while an eQTL mapped to a locus other than that of the encoding ORF suggests that its expression is regulated by *trans*-acting regulatory factors. The eQTL approach has been deployed to dissect the genetic basis of freezing tolerance in wheat at a major QTL on chr. arm 5AL (Vagujfalvi et al. 2003, 2005; Galiba et al. 2005). Due to the high cost for profiling RNA samples of an entire mapping population, transcriptome profiling based on microarrays is better suited for studies involving a limited number of samples extracted from congenic strains differing at key genomic regions (e.g. near-isogenic lines; Guo et al. 2007; Jukanti et al. 2008); alternatively, bulked samples obtained from the tails of mapping populations could be used.

The interpretation of the results obtained from profiling experiments carried out under controlled conditions should critically consider the experimental conditions utilised to mimic the abiotic stress, an issue particularly important for studies targeting tolerance to drought. In several studies aimed at identifying genes differentially expressed under drought, plants or parts thereof were exposed to severe stress intensity in a very short time, commonly only a few hours (reviewed in Hazen et al. 2003). These experimental conditions will be more damaging at the cellular level if compared to similar levels of water deficit that plant tissues experience in the field, where dehydration unfolds over a

prolonged period of time (commonly days or weeks), thus allowing for a more proper activation of the molecular mechanisms leading to those beneficial adaptive responses (e.g. osmotic adjustment, early flowering, thickening of leaf cuticles, etc.) able to counteract the negative effects of drought. In barley, the changes in gene expression observed in excised leaves following a rapidly induced dehydration were poorly correlated (from 0.19 to 0.41, according to the genotype and the intensity of dehydration) with those attained under a slower dehydration regime (in pots) which mimicked more closely a drought episode in field conditions (Talamè et al. 2007). Therefore, molecular results obtained under artificial conditions mimicking a water deficit should be dealt cautiously and duly validated prior to their utilisation in a more applicative context (genetic engineering and/or MAS).

The sequencing of entire genomes has highlighted our rudimentary understanding of gene function, since a large portion of ORFs have not been assigned to a known function (Pop and Salzberg 2008). Additional clues to gene function and the changes caused by abiotic stress can be acquired through profiling of the proteome (Bahrmann et al. 2004, 2005; Hajheidari et al. 2007; Khan et al. 2007; Song et al. 2007; Zhu et al. 2007) and metabolome (Morgenthal et al. 2006; Roessner et al. 2006) which, as compared to the transcriptome, are functionally “closer” to phenotypic traits (Sakurai and Shibata 2006) and thus can also account for the effects due to post-transcriptional and post-translational regulation (e.g. protein phosphorylation). Thus, within the life sciences there is a need for determination of the biological function of these so-called orphan genes, some of which may represent valuable targets for direct manipulation. Nonetheless, it should be appreciated that both proteomics and metabolomics can report changes occurring in only a rather limited portion of the genome; additionally, proteomics is often unable to detect the changes in gene products (e.g. transcription factors) that, despite their low level, likely play an important role during the early phase of the plant response to adverse conditions (Shinozaki and Yamaguchi-Shinozaki 2007). Profiling the proteome of a mapping population offers the opportunity to identify QTLs influencing protein quantity (PQLs, Protein Quantity Loci; Damerval et al. 1994; Consoli et al. 2002). Co-localisation of a PQL with the protein-coding locus would indicate that allelic differences at that locus influence the final level of the protein, whereas co-localisation between a PQL and a QTL for a different trait would suggest a possible association between the candidate protein and trait variation (for an example on freezing tolerance in wheat, see Vagujfalvi et al. 2003).

In wheat and barley, transcriptome profiling (see Table 18.2) has targeted changes in gene expression caused by salinity (Ozturk et al. 2002; Atienza et al. 2004; Ueda et al. 2004), dehydration (Atienza et al. 2004; Way et al. 2005; Tondelli et al. 2006; Xue et al. 2006; Mohammadi et al. 2007; Talamè et al. 2007), extreme temperatures (Faccioli et al. 2002; Atienza et al. 2004; Gulick et al. 2005; Tondelli et al. 2006; Oshino et al. 2007), toxic minerals (Atienza et al. 2004; Furukawa et al. 2007) and nutrients (Jukanti et al. 2008; Ruuska et al.

2008). As an example, Tondelli et al. (2006) searched for candidate genes for cold- and drought-stress tolerance. In particular, 13 genes encoding for transcription factors and upstream regulators were investigated on the parental genotypes, and were mapped on three barley mapping populations, thus allowing for the assembly of a consensus function map reporting the position of barley QTLs for cold and drought tolerance. This comparative analysis showed that a cluster of six *HvCBF* genes co-mapped with the *Fr-H2* cold tolerance QTL, while no QTLs for the same trait was assigned to chr. 7H which harbours *ICE1* and *FRY*, two putative barley regulators of CBF expression, thus suggesting CBF gene, rather than its two regulators, as best candidates for cold tolerance. Of the 12 drought tolerance QTLs of the consensus map, four were associated with regulatory candidate genes, on chrs. 2H, 5H and 7H, and two with effector genes located on chrs. 5H and 6H.

The wealth of stress-related genes evidenced by transcriptomics studies and the plethora of QTLs so far reported in the literature, inevitably implies their co-location by chance alone. Therefore, the circumstantial evidence provided by transcriptome studies should be treated with due caution, particularly when establishing a more solid cause-effect relationship between the variation observed in gene expression and the variation in the target trait. A number of forward- and reverse-genetics approaches allow us to further investigate such relationships. Among these, TILLING and EcoTILLING provide valuable opportunities to elucidate the functional role of target sequences (Till et al. 2007). In the Triticeae, the collection of mutagenized materials that have been assembled in barley (Caldwell et al. 2004; Talamè et al. 2008) and wheat (Slade et al. 2005; Slade and Knauf 2005) for TILLING purposes can also be exploited in a forward-genetics fashion to identify mutants for target traits known to influence tolerance to abiotic stresses. As an example, a collection of ca. 50 root mutants selected from a forward screening of the sodium-azide treated M₄ progeny derived from cv. Morex (Talamè et al. 2008) is being characterised with the objective to map and eventually clone the most interesting mutants (Talamè et al. unpublished).

18.3 QTLs and Genes for Tolerance to Abiotic Stress

On a global scale, abiotic stress is the major cause of yield losses in the Triticeae. Because wheat and barley are prevalingly grown under rainfed conditions and often in marginal areas characterised by low and unpredictable rainfall, and the concomitant presence of other constraints (e.g. toxic ions that impair root growth), drought is the most devastating abiotic stress for these two crops, as shown by the drastic reduction in barley and wheat yield caused by the prolonged drought spell in Australia in 2006 and 2007 (<http://www.lilith-ezine.com/articles/environmental/Australian-Drought.html>). Due to its complexity and poor heritability,

drought tolerance has often been considered one of the most recalcitrant traits to improve, hence representing a prime candidate and target for genomics-assisted improvement.

18.3.1 Tolerance to Drought

By definition, a crop undergoes a drought stress when the amount of water available either through rainfall and/or irrigation is insufficient to meet its evapo-transpirative demand. The more prolonged this time-period is, the more drastic the negative consequences on yield. Another crucial factor is the timing of the drought episode in relation to the vegetative or reproductive stage of the crop (Richards 1996, 2006; Araus et al. 2002; Ugarte et al. 2007; Fan et al. 2008). Additionally, drought episodes are erratic, with unpredictable distributions throughout years and seasons. Under this respect, a drought episode is most harmful at flowering, when the reproductive fertility can be drastically curtailed by an insufficient supply of photosynthates and changes in phyto-regulators (e.g. excessive accumulation of abscisic acid; Saini and Westgate 2000; Fresneau et al. 2007). Moreover, environments with Mediterranean climates are frequently subjected to terminal drought episodes during the grain-filling stage. For these cases, the amount of stem reserves and the relocation capacity play a key role to cope with the associated grain weight reduction (Salem et al. 2007; Ehdaie et al. 2008).

For categorizing the mechanisms conferring tolerance to drought, we have adopted the nomenclature suggested by Levitt (1972), whereby the features that allow the plant to escape drought (e.g. early flowering in crops grown in Mediterranean-like environments) are distinct from those influencing resistance to drought, with the latter further categorised in terms of dehydration avoidance and dehydration tolerance. Dehydration avoidance depends on maintenance of turgor through an improved water balance, either through increased water uptake (e.g. deeper roots, osmotic adjustment) and/or reduced water loss (e.g. increased leaf waxiness). Dehydration tolerance involves biochemical mechanisms (e.g. accumulation of compatible solutes to preserve membrane integrity, water-soluble carbohydrates relocation, etc.) that enable the cell to tolerate the negative effects caused by cellular dehydration.

The limited success in improving drought resistance through molecular approaches is primarily due to the difficulty in identifying and accurately measuring the key physiological determinants of yield under drought conditions in the field (Blum 1988, 2005; Ludlow and Muchow 1990; Turner 1997; Passioura 2002; Nguyen and Blum 2004; Tuberosa 2004; Passioura 2007; Cattivelli et al. 2008). For example, root meristems capable of adjusting osmotically at a given water potential will positively impact root mass and final yield to the extent that deeper roots will improve the extraction of additional moisture from the soil. However, if additional moisture is unavailable at deeper soil

layers, a condition quite common in many drought-prone environments (e.g. the Mediterranean Basin), growing larger/deeper roots will not be advantageous and might even negatively influence final yield due to excessive partitioning of photosynthates to the roots and the high metabolic cost required for sustaining the functions of a large root system. Under water-stressed conditions, a critical factor in determining the effects on yield due to variation in physiological traits is the water status of the plants. Therefore, when evaluating physiological traits for their association with grain yield, it is important to take into due consideration the water status of the plant from where samples are collected. An example is provided by the concentration of abscisic acid (ABA), a phytohormone involved in the adaptive response of the plant to decreased soil moisture and other stresses (Quarrie 1991; Sharp et al. 2004; Xiong 2007). In the Triticeae, although early work reported QTLs for the accumulation of ABA in dehydrated leaf tissue in both barley (Sanguineti et al. 1994) and wheat (Quarrie et al. 1994b), additional evidence on the role of the QTLs for ABA accumulation in regulating field performance under drought-stressed conditions has not been reported. Finally, a thorough evaluation of the effects of a target locus should duly consider the influence of management factors such as planting density in the field, a feature which will greatly influence crop evapotranspiration.

18.3.1.1 Barley

In barley, the most extensive QTL search under conditions of water deficit has involved a mapping population of recombinant inbred lines (RILs) derived from Tadmor (drought tolerant) \times Er/Apm (adapted only to specific dry environments) that was tested in a number of experiments conducted under controlled conditions and in the field (Table 18.1A). Teulat et al. (1998) searched for QTLs for osmotic adjustment (OA) at an early vegetative stage using two water treatments corresponding to soil moisture of 14 and 100% of field capacity. Two QTLs on chrs. 7H and 6H controlled variation in several OA-related traits. The chr. 7H region near *Ac13*, involved in osmoregulation, is conserved in wheat (Morgan 1991) and rice (Lilley 1996). When the same set of materials was analysed for OA and water-soluble carbohydrate (WSC) concentration under similar conditions and with a more complete map, eight additional regions comprising 22 QTLs were identified, thus increasing to 13 the total number of chromosomal regions (32 QTLs) controlling traits related to plant water status and/or OA in this particular genetic background (Teulat et al. 2001a). Interestingly, the effects of the QTL on chr. 6H in regulating RWC was confirmed when the RILs were evaluated in several rainfed Mediterranean sites (Teulat et al. 2001b, 2003). The RILs derived from Tadmor \times Er/Apm were also investigated for QTLs for carbon isotope discrimination ($\delta^{13}\text{C}$), a trait that has been associated with water-use efficiency (WUE) and yield stability in drought-prone environments (Teulat et al. 2002; Araus et al. 2002; Condon et al. 2004; Tambussi et al. 2007; Xue et al. 2007). QTLs for grain $\delta^{13}\text{C}$ were

detected in experiments carried out in three Mediterranean field environments differing in water availability (Teulat et al. 2002). Eight QTLs for $\delta^{13}\text{C}$ were found to co-locate with QTLs for physiological traits related to plant water status or OA, and/or for agronomic traits previously measured on the same population. QTLs for $\delta^{13}\text{C}$ of the shoot tissue have been reported also in Derkado \times B83-12/21/5, with one of the identified QTLs also influencing stem $\delta^{13}\text{C}$ (Ellis et al. 2002).

A number of differentially expressed sequence tags (dESTs) and candidate genes for drought response identified by Ozturk et al. (2002) were mapped in Tadmor \times Er/Apm (Diab et al. 2004). The survey of 100 sequenced probes from two cDNA libraries previously constructed from drought-stressed barley (Ozturk et al. 2002) and 12 candidate genes allowed for the addition of 33 loci to a previously published map (Diab et al. 2004). Wild barley (*Hordeum vulgare* ssp. *spontaneum*) is a valuable source of alleles for resistance to abiotic stresses (Forster et al. 2000; Ivandic et al. 2000; Robinson et al. 2000; Turpeinen et al. 2001). A population of 123 BC₁DH lines derived from a *H. vulgare* (cv. Barke) \times *H. spontaneum* cross were investigated in three Mediterranean countries to identify agronomically favourable QTL alleles contributed by the wild parent (Talamè et al. 2004). Among the 81 putative QTLs found to influence growth habit, heading date, plant height, ear length, ear extrusion, grain yield and/or 1,000-grain weight, in 43 cases (53%) the wild parental line contributed the alleles with favourable effects. As for grain yield, although the majority (65%) of the favourable QTL alleles was inherited from *H. vulgare*, at six QTLs the alleles increasing grain yield were contributed by *H. spontaneum*. Interestingly, these QTLs only marginally affected heading date, thus indicating that escape did not play an important role in determining yield. These results indicate that the advanced-backcross QTL approach (ABQA) provides a useful germplasm enhancement strategy for identifying wild progenitor QTL alleles capable of improving grain yield of the related crop under stressful conditions. An ABQA involving *H. spontaneum* was also exploited by Forster et al. (2004). Similarly to what was reported by Talamè et al. (2004), also in this case the wild parental line contributed a number of positive alleles for yield, with major effects due to QTLs that clustered around major genes controlling flowering time (*sgl1*), plant stature (*sdw1* and *arie.GP*) and ear type (*vrs1*). Based on these results, Forster et al. (2004) postulated that MAS for these morpho-physiological features can be used to match crop phenology to field environment. The evaluation under Mediterranean conditions of 194 RILs from an *H. vulgare* (Arta) \times *H. spontaneum* cross showed that the most important effects under drought stress were due to QTLs for plant height on chrs. 2H, 3H and 7H. The QTLs for plant height, particularly the one on 3H, showed concurrent effects on heading date, grain yield and biological yield (Baum et al. 2003). Under rainfed conditions in some Mediterranean countries, plant height is a very important trait, since straw can still provide a meaningful source of feed for grazing animals in dry years.

18.3.1.2 Wheat

Different morpho-physiological traits have been reported to account for the variability observed in yield among RILs of mapping populations (Table 18.1A). As an escape mechanism, flowering time plays a major role in Mediterranean countries and in parts of Australia characterised by a Mediterranean-like climate where summer droughts prevail after flowering (e.g. Western Australia). Under these conditions, several studies have reported QTLs characterised by a negative association between heading date and grain yield (Araus et al. 2002; Maccaferri et al. 2008). Among the traits contributing to an improved performance under water-limited conditions, early vigor and root architecture are receiving increasing attention (Kimurto et al. 2005). It has been speculated that a moderate soil drying could accelerate the elongation of some lateral roots (Ito et al. 2006). Variations in length of first-order lateral roots and their responses to soil drying are positive features for water absorption when resources are limited. In wheat, both seminal and nodal roots remain functionally active throughout the life of the plant. In this respect, seminal roots may be equally or even more important than nodal roots for sustaining yield because they can grow deeper in the soil profile (Manske et al. 2002). Moreover, under conditions of limited soil moisture, nodal roots do not develop or their growth is limited; therefore, in these cases, plants reach maturity relying prevalently on their seminal roots. The role of root architecture on WUE and agronomic performance in pots and in the field under WW- and WS-conditions was investigated by Ehdaie et al. (2003) using the centric translocations of the short arm of chr. 1R of rye in bread wheat. The 1RS translocations in Pavon delayed maturity, reduced plant height in some cases, and increased root biomass. Root biomass and grain yield were positively associated under both WS- and WW-field conditions. The 1RS translocations increased grain yield and grain weight, especially under WW-field conditions. The 1RS translocations, in general, were more tolerant to field environmental stresses than Pavon, thus suggesting the possibility of deploying 1RS.1AL and 1RS.1DL translocations in wheat breeding programs. Kubo et al. (2007) evaluated traits related to root penetration (RP) ability in Jennah Khetifa \times Cham1 using paraffin-Vaseline discs to mechanically impede root penetration. Two closely-linked markers, *Xgwm617a* and *Xgwm427b*, on chr. arm 6AL were associated with the number of roots penetrating the disc and root penetration index. For root biomass, a QTL was detected near *Xgwm11* on chr. 1B. The absence of co-located QTLs suggested that root penetration ability was controlled separately from total number of seminal and crown roots and root biomass. Association mapping was applied to a set of 57 elite durum wheat cultivars to search for QTLs able to affect root architecture at the seedling stage (Sanguineti et al. 2007). These cultivars were sampled from a larger panel of durum accessions characterised by a high LD (Maccaferri et al. 2005). In total, 15 chromosome regions showed significant effects on one or more root architectural features. A number of these regions also influenced shoot traits and, in some cases, plant height measured in field

conditions. Major effects were detected on chr. arms 2AL (at *Xgwm294*), 7AL (at *Xcfa2257* and *Xgwm332*) and 7BL (at *Xgwm577* and *Xcfa2040*).

Remobilisation of WSC from the stem and leaves has been recognised as a mechanism able to mitigate the negative effects of post-anthesis drought tolerance (Blum 1988; Araus et al. 2002; Reynolds et al. 2007a). In bread wheat, QTLs for stem-reserve mobilisation were mapped on chrs. 2D, 5D and 7D of the ITMI population derived from W7984 × Opatá 85 (Salem et al. 2007). In a DH population derived from two Chinese common wheat cultivars and tested under WW- and WS-conditions, Yang D.L. et al. (2007a) reported 48 additive QTLs for nine traits related to stem WSC accumulation and remobilisation at grain-filling period under drought-stressed conditions. Additionally, 62 epistatic pairs of QTLs were described. Most QTLs showed significant interactions with environments (QEIs), whose contributions were two- to four-fold higher than those of their corresponding QTLs. QTLs for leaf senescence and WSC remobilisation were also reported by Snape et al. (2007) in four DH populations tested across WW- and WS-conditions. Rebetzke et al. (2008) phenotyped three wheat mapping populations (Cranbrook × Halberd, Sunco × Tasman and CD87 × Katepwa) for WSC and other agronomic traits across multiple environments. The range for WSC concentration (WSC-C) was large among progeny and genetic control complex, with 7–16 QTLs being identified for WSC-C in each population. Heritability was lower ($h^2 = 0.32–0.54$) for WSC mass per unit area (WSC-A) and fewer significant QTLs (4–8) were identified in each population, while sizes of individual genetic effects varied between populations but were repeatable across environments. Genotypes with high WSC-C were commonly shorter, flowered earlier and produced significantly fewer tillers than those of low WSC-C. This resulted in similar yields, lower final biomass and fewer grains per m², but greater dry weight partitioning to grain, kernel weight and less grain screenings in high compared with low WSC-C genotypes. By contrast, lines high for WSC-A produced more fertile tillers associated with similar or greater anthesis and maturity biomass, grain number and yield, yet similar kernel weight or size compared with genotypes with low WSC-A thus supporting an important role for WSC-A in assuring stable yield and grain size. Nonetheless, the small effects of many independent WSC QTLs may limit their direct use for marker-aided selection in breeding programs (Rebetzke et al. 2008).

Parameters of chlorophyll fluorescence kinetics (PCFKs: e.g. Fv, Fm, Fv/Fm, Fv/Fo) were measured by Yang D.L. et al. (2007b) to investigate dehydration tolerance in wheat. The same DH population used to study WSC was also used to analyse the correlation between PCFKs and chlorophyll content (ChlC) and to map QTLs at the grain-filling stage under conditions of both WS- and WW-conditions. A total of 14 additive QTLs (nine and five QTLs detected under WS- and WW-conditions, respectively) and 25 pairs of epistatic QTLs for PCFKs were mapped on chrs. 6A, 7A, 1B, 3B, 4B and 7D. Several QTL clusters that were detected on chrs. 1B, 7A and 7D were identified under only one water regime, thus suggesting an adaptive role.

In barley, QTLs for PCFKs were investigated in 194 RILs derived from an Arta \times *H. spontaneum* cross (Guo et al. 2008). A total of nine and five QTLs were identified under well-watered and water-stressed conditions at post-flowering stage. No common QTL was detected except one for chlorophyll content, which was identified in both growth conditions. The two QTLs with the largest effect on Fv/Fm were identified on chr. 2H.

In durum wheat, Maccaferri et al. (2008) searched for QTLs for grain yield, heading date and plant height in 249 RILs derived from Kofa \times Svevo that was evaluated in ten rainfed and six irrigated environments characterised by a broad range of grain yield (from 5.6 to 58.8 q ha⁻¹). Among the 16 QTLs that affected grain yield, two major QTLs on chr. arms 2BL and 3BS showed significant effects in eight and seven environments, respectively. In both cases, extensive overlap was observed between the LOD profiles of grain yield and plant height, but not with those for heading date, thus indicating that escape did not play an important role in determining yield in this population. For both plant height and grain yield, notable epistasis between chr. arms 2BL and 3BS QTLs was detected across several environments, with the parental combinations providing the higher performance. These two QTLs exhibited significant additive and epistatic effects also on ear peduncle length and kernel weight (Maccaferri et al. unpublished results).

18.3.2 Tolerance to Salinity

Excessive salinity afflicts ca. 20% of the cultivated surface and 33% of irrigated agricultural lands world-wide (Flowers and Yeo 1995; Munns 2005; Tuteja 2007; Witcombe et al. 2008), while ca. 1% of arable land is lost annually due to excessive salinization. At the crop level, experimental evidence indicates the quantitative nature of salt tolerance and the influence of the environment (Blum 1988; Rao and McNeilly 1999; Ashraf 2002; Munns and Richards 2007). Tolerance to salinity can be achieved by controlling the rate of salt accumulation so that cytoplasmic concentrations do not become toxic while maintaining the supply of carbon to sinks (Munns 1993; Munns 2005; Munns et al. 2006; Reynolds et al. 2005). Nonetheless, salt tolerance has not always been associated with lower accumulation of Na⁺ in the shoot, as shown by the lack of correlation between shoot Na⁺ accumulation and salt tolerance reported in bread wheat, a finding which prompted Genc et al. (2007) to suggest tissue tolerance as a better selection criterion for salinity tolerance in wheat. Tissue tolerance in leaves could be identified by mapping loci that control salt tolerance but not shoot Na⁺ accumulation (Collins et al. 2008). In these cases, tolerance to salinity is likely achieved by sequestering toxic Na⁺ in the vacuoles or, alternatively, in starch granules in parenchyma cells in the stele (Munns 2007).

In barley, genetic differences in tolerance to saline conditions have been reported (Richards et al. 1987; Slavich et al. 1990) and a number of genes seemingly affecting Na^+ influx and concentration in cells have been described (Schachtman and Liu 1999). QTLs affecting salt tolerance at germination and seedling stage were reported in two barley populations (Steptoe \times Morex and Harrington \times TR306; Mano and Takeda 1997). In both cases, the most important QTLs occurred at different loci on chr. 5H. These QTLs differed from those controlling salt tolerance at germination, suggesting that salt tolerance at germination and at the seedling stage were controlled by different loci. A DH population derived from the cross between cv. Derkado and B83-12/21/5, a *H. spontaneum* accession, tested under saline solution in hydroponics evidenced major effects on seedling growth near the *ari-eGP* (on chr. 5H) and the *sdw1* (on chr. 3H) dwarfing genes, which were also found to influence C- and N-isotope discrimination (Ellis et al. 2002).

In durum wheat, QTL mapping carried out by Munns and coworkers led to the identification of *Nax1* (chr. arm 2AL) and *Nax2* (chr. arm 5AL), two loci controlling shoot Na accumulation (Munns et al. 2003; Davenport et al. 2005; James et al. 2006). At both loci, the alleles for Na^+ exclusion derived from introgressions of a *Triticum monococcum* accession. Importantly, marker-assisted backcrossing (MABC) has allowed for the introgression of the Na -exclusion genes into five cultivars of bread wheat by using durum \times bread wheat inter-specific crosses; the resulting lines are currently under field evaluation in Australia (Colmer et al. 2006; R. James and R. Munns, personal communication). *TmHKT7-A2* and *TmHKT1;5-A* (*HKT8*) are the candidates for *Nax1* and *Nax2*, respectively (Huang et al. 2006; Byrt et al. 2007), while *TaHKT1;5-D*, the corresponding homologue in the D genome, has been suggested as the candidate for the *Kna1* gene on chr. arm 4DL which accounts for the greater salt tolerance of bread wheat compared to durum wheat (Byrt et al. 2007). Interestingly, *OsHKT1;5*, the apparent orthologue of the *TaHKT1;5* genes in rice, affects Na^+ accumulation and underlines a major QTL for salt tolerance on rice chr. 1 (Ren et al. 2005). HKTs transporters have been shown to lower shoot Na^+ accumulation by facilitating unloading of Na from the xylem (Ren et al. 2005; Sunarpi et al. 2005). In Arabidopsis, an *HKT1* homologue (*AtHKT1*) underlines a natural QTL (Rus et al. 2006), and its contribution to shoot Na^+ exclusion and salinity tolerance in this species had been demonstrated by insertion mutagenesis (Sunarpi et al. 2005). Therefore, ample experimental evidence underlines the importance of the HKTs transporters in the regulation of salt tolerance across a broad range of plant species.

18.3.3 Tolerance to Low Nutrients

Tolerance to low nutrients is of utmost importance in countries where farmers do not have access to the economic resources required to purchase chemical

fertilizers. Additionally, increasing tolerance to nutrient-deficiency of cereals represents a more cost-effective and long-term sustainable solution than relying on fertilizer application. The sharp increase in the cost of oil has also made N fertilizer considerably more expensive. Phosphorus (P) deficiency can also be a major factor limiting crop yield, especially in low-input agricultural systems. Unlike N fertilizers, P fertilizers cannot be synthesized artificially thus representing a more critical problem in the long-term, particularly considering that P reserves have been forecast to become exhausted by the end of this century. Additionally, leaching of P and N into surface and sea water also causes widespread environmental problems. Clearly, improving N- and P-use efficiency of Triticeae crops would allow for a more profitable and sustainable agriculture.

18.3.3.1 Nitrogen

The release of cultivars able to attain acceptable yield levels even under low-N supply can be achieved by improving N-use efficiency (NUE), a complex trait involving both the uptake and allocation of N to the grains, with the latter largely influenced by the capacity to remobilise N from the vegetative organs to the seeds. Therefore, N remobilisation is crucial for yield in the Triticeae, particularly under conditions of post-anthesis stress. While chloroplasts contain most of the reduced N present in photosynthetically active leaf cells, the main pathways involved in the degradation of their proteins prior to the reallocation of the resulting amino acids are largely unknown. An example of the complexity entailed by the study of a “molecular” trait related to N metabolism is provided by GS1, an enzyme that appears to play a key role in N metabolism and leaf senescence, a change usually hastened by a number of abiotic stresses. GS1 is the product of multiple genes with complex promoters that drive the expression of the genes in an organ- and tissue-specific manner and in response to a number of environmental signals affecting the nutritional status of the cell. GS activity is also regulated post-translationally in a manner that involves 14-3-3 proteins and phosphorylation. GS and plant N metabolism is best viewed as a complex matrix continually changing during the development cycle of plants (Mifflin and Habash 2002).

In barley, a population (146 RILs) derived from a cross between two varieties differing markedly in grain protein concentration was evaluated to identify QTLs for N uptake, storage and remobilisation in flag leaves and for developmental parameters and grain protein (GP) accumulation (Mickelson et al. 2003). The most noticeable overlap of QTL profiles for the investigated traits occurred on chrs. 3H and 6H, where alleles with lower efficiency for N remobilisation were associated with lower yield and higher levels of total or soluble organic N during grain filling thus suggesting that genes directly involved in N recycling or genes regulating N recycling may be located on these chromosomes. However, the most important QTL for GP concentration (chr. 6H) while overlapping with QTLs for nitrate and soluble organic N (with a negative association between the effects on GP concentration and leaf soluble N levels

during grain filling) did not co-localize with QTL for N remobilisation. The same mapping population was also investigated for leaf amino-, carboxy- and endopeptidase activities relative to previously reported QTLs for GP, leaf-N storage and remobilisation (Yang et al. 2004). The results suggested that major endopeptidases are not instrumental in leaf-N remobilisation or the control of grain protein accumulation. Additionally, QTL profiles for aminopeptidases when compared to those for N remobilisation indicated no functional role for the enzymes assayed in plant-N recycling. Conversely, the QTL data suggested that one or several carboxypeptidase isoenzymes may contribute to N recycling.

In hexaploid wheat, the role of root in N uptake was investigated by An et al. (2006) using 120 DH lines derived from Hanxuan 10 \times Lumai 14 that were tested in the field and in hydroponics in conditions of low N (LN) and high N (HN) supply. Shoot and root dry weight (SDW and RDW), tiller number (TN) and NUP (total plant N uptake) investigated in hydroponics were significantly and positively correlated with NUP under both LN- and HN-conditions in the field trials. Nine and eight QTLs for NUP were detected under LN- and HN-conditions in the field trials, respectively. Among the four to five QTLs for SDW, RDW, TN and NUP detected in hydroponics, one for SDW, three for RDW and two for TN overlapped with QTLs for NUP under LN- or HN-conditions in the field. These overlaps occurred on chrs. 1B, 2DL and 6AL. The positive correlation found at these overlapping QTLs between the effects in hydroponics and NUP in the field suggested that seedling root and shoot vigor is an important factor influencing N uptake in hexaploid wheat. In durum wheat, QTLs for root architecture and biomass at an early growth stage have been identified by Sanguineti et al. (2007) by means of an association mapping approach.

In a bread wheat mapping population derived from the cross Chinese Spring \times SQ1, QTL clusters for Glutamine Synthetase (GS) activity on chrs. 2A and 4A coincided with the location of a GS and GSr gene, respectively. Although QTL alleles for higher GS activity at these loci were associated with higher grain N, they showed little or no effects on grain yield components (Habash et al. 2007). QTL effects for peduncle N were positively correlated with those for grain- and flag leaf-N assimilatory traits, suggesting that stem N can be predictive of grain N status in wheat. A major QTL effect for ear number per plant was detected on chr. 6B which was negatively associated with leaf fresh weight, peduncle N, grain N and grain yield, leading Habash et al. (2007) to suggest that this QTL is involved in processes defining the control of tiller number and assimilate partitioning. Using both a conceptual model of C/N plant functioning and a root architecture description, Laperche et al. (2006) searched for QTLs controlling traits associated with N uptake and NUE using a population of 120 DH lines derived from the cross between two cultivars contrasted for tolerance to N deprivation and grown in hydroponic rhizotrons under N-limited conditions. Among the 32 QTLs detected, six were associated with root traits, six with model efficiencies and 20 with nitrogen-integrative traits. Based on these results, Laperche et al. (2006) suggested that coupling

QTL data with a functioning model and a root architecture description could serve to identify suitable ideotypes. DH lines (222) of the same cross tested in seven environments under both high- and low-N supplies revealed major influences of *Ppd-D1* (locus for photoperiod sensitivity on chr. 2D), *Rht-B1* (locus for plant height on chr. 4B) and *B1* (locus for the presence of awns on chr. 5A) on the 233 N-related QTLs that were detected (Laperche et al. 2007). Important specific QTLs were also detected on chrs. 3D, 4B, 5A and 7B. The genome-wide transcription profiling study carried out by Ruuska et al. (2008) investigated the molecular basis of adaptation to N limitation of wheat as related to the accumulation of carbon (C) stem reserves for later remobilisation to grain. Under N-deprived conditions, fructan accumulation in stems was accompanied by elevated transcripts for a number of fructosyltransferases (FTs) and fructan 6-exohydrolase (6-FEH) genes. Clustering analysis identified a grouping that included several FTs and a number of genes thought to be involved in regulation of storage C metabolism or senescence in other systems (Ruuska et al. 2008). One problem frequently encountered when investigating the expression of multigene families is the identification of unique sequence polymorphism that would enable their mapping. An interesting example is provided by the study on nitrate reductase, nitrite reductase, glutamate dehydrogenase and glutamate synthase (GOGAT) in *Triticum aestivum*, *T. durum*, *T. monococcum*, *T. speltoides* and *T. tauschii* (Boisson et al. 2005). In the ca. 8 kb of gene sequences explored, the polymorphism rate was higher for non-coding regions, where it ranged from 1/60 to 1/23, than for coding regions (from 1/110 to 1/40) except when the hexaploid D genome was compared to that of *T. tauschii* (1/800 and 1/816, respectively). Although the assembly of specific primer pairs is time-consuming and expensive because of the sequencing, Boisson et al. (2005) demonstrated that polymorphism detection enables one to discriminate heterologous, homeologous and even paralogous copies.

18.3.3.2 Phosphorus

Because of the poor mobility of P in the soil, root architecture plays an important role in P uptake (Desnos 2008). A population of 92 DH lines derived from P-deficiency tolerant wheat variety Lovrin 10 and P-deficiency sensitive variety Chinese Spring was tested in pots at the seedling stage under contrasting P conditions (Su et al. 2006). Several of the 20 and 19 QTLs detected under low- and high-P condition, respectively, clustered in three regions near *Xgwm251* (chr. 4B), *Xgwm271.2* (chr. 5A) and *Xgwm121* (chr. 5D). Notably, two of the above QTLs were closely linked with vernalization requirement genes *VRN-A1* (on chr. 5A) and *VRN-D1* (on chr. 5D), thus once more underlining the importance of phenology in determining the adaptive response to abiotic stress.

Because plants respond to P deficiency through multiple strategies, including the synthesis of high-affinity P transporters, the expression pattern of *TaPT2*, a putative wheat high-affinity phosphate transporter, was examined in wheat roots and leaves under P-deficient conditions (Tittarelli et al. 2007). The results

showed that a 579-bp fragment of the *TaPT2* promoter is sufficient to drive the expression of the GUS reporter gene specifically in roots of P-deprived wheat. Additionally, this *TaPT2* promoter fragment was also able to drive expression of the GUS reporter gene in transgenic *Arabidopsis thaliana* grown in P-deprived conditions. An *in silico* comparative analysis showed the presence of conserved regions and candidate regulatory motifs between this wheat promoter and P-transporter promoters from barley, rice and *Arabidopsis*, a result that suggests the presence of conserved *cis*-acting elements and *trans*-acting factors able to regulate the *TaPT2* promoter in a tissue-specific and P-dependent fashion in both monocots and dicots.

Huang et al. (2008) undertook a metabolomic study to unravel the metabolic processes for adaptation to low P conditions. Their results showed that mildly P-deficient plants accumulated di- and trisaccharides (sucrose, maltose, raffinose and 6-kestose), especially in shoots. Severe P deficiency increased the levels of metabolites related to ammonium metabolism in addition to di- and trisaccharides, but reduced the levels of phosphorylated intermediates (e.g. glucose-6-P, fructose-6-P, inositol-1-P and glycerol-3-P) and organic acids (e.g. alpha-ketoglutarate, succinate, fumarate and malate). The results revealed that P-deficient plants modify carbohydrate metabolism initially to reduce P consumption, and salvage P from small P-containing metabolites when P deficiency is severe, which consequently reduced levels of organic acids in the tricarboxylic acid (TCA) cycle. Additionally, a sharp increase in glutamine and asparagine was observed in both shoots and roots of severely P-deficient plants. Based on these results, Huang et al. (2008) proposed a strategy for improving tolerance to low P based on an altered partitioning of carbohydrates into organic acids and amino acids to enable more efficient utilisation of carbon in P-deficient plants.

18.3.4 Tolerance to Aluminium Toxicity

Aluminium (Al) toxicity is one of the major constraints for plant development in acid soils in many countries (e.g. Brazil). At low pH, the phytotoxic Al^{3+} cation is released into the soil solution where it inhibits root growth, hence reducing the ability of plants to acquire water and nutrients (Ma et al. 2004). Among the Triticeae, rye shows the highest tolerance to Al. In rye, at least four independent and dominant loci, *Alt1*, *Alt2*, *Alt3* and *Alt4*, located on chr. arms 6RS, 3RS, 4RL and 7RS, respectively, provide one of the most efficient group of genes for Al tolerance (Miftahudin et al. 2002, 2005; Matos et al. 2007).

In hexaploid wheat, Milla and Gustafson (2001) evaluated a RIL population (BH1146 \times Anahuac) segregating for Al tolerance and a series of deletion lines to obtain a more detailed genetic linkage map of the chr. arm 4DL, which previous studies had shown to harbour genes conferring resistance to Al. The *Alt_{BH}* gene was confined to a 5.9-cM interval between markers *Xgdm125* and *Xpsr914*. The results obtained with the wheat deletion lines for chr. arm 4DL

physically mapped the *Alt_{BH}* gene to the distal region of the chromosome where the ratio between the genetic/physical distances appears feasible for the map-based cloning of the gene (Milla and Gustafson 2001).

The comparative analysis of major loci and QTLs for Al tolerance reported in rice (Xue et al. 2007), sorghum (Kochian et al. 2005; Magalhaes et al. 2007), barley (Wang et al. 2007), wheat (Raman et al. 2005), oat (Wight et al. 2006) and rye (Matos et al. 2005), suggests the existence of possible homologies between Al-tolerance loci in cereals (Collins et al. 2008). As an example, *Alt3* in rye is tightly linked to markers mapping near *Alt_{BH}* in wheat, suggesting a possible homology relationship for these two loci (Matos et al. 2005).

The most common mechanism of Al tolerance relates to the Al-activated extrusion of Al-chelating anions (e.g. malate and citrate) from the root tips, followed by the formation of non-toxic Al complexes in the apoplast or rhizosphere (Ma et al. 2004; Delhaize et al. 2007). In wheat, the *ALMT1* gene encodes a malate transporter that has been associated with malate efflux and Al tolerance (Sasaki et al. 2004). Transgenic barley plants expressing *ALMT1* were generated to assess their ability to exude malate and withstand Al stress (Delhaize et al. 2004). The transgenic barley showed a high level of Al tolerance when grown in both hydroponic culture and on acid soils. These findings confirm that *ALMT1* is a major Al-tolerance gene and demonstrate the feasibility to improve tolerance to acid soils through a transgenic approach.

In common wheat, Raman et al. (2006) identified molecular markers targeting insertion/deletion (indel) and SSR repeats within the intron three region of the *ALMT1* gene. Subsequently, allelic diversity was investigated more extensively with four markers specific to intron three, exon four and promoter regions of *ALMT1* (Raman et al. 2008). When these markers were compared in 179 cultivars from international wheat breeding programs that were released during the last century, six different promoter haplotypes were identified. A previous study showed that Al resistance in common wheat was not associated with a particular coding allele for *ALMT1* but was correlated with blocks of repeated sequence upstream of the coding sequence (Sasaki et al. 2006). Analysis of a population of 278 landraces and subspecies of wheat revealed several new promoter haplotypes and showed that most of the promoter alleles associated with Al resistance already existed in Europe, the Middle East and Asia prior to dispersal of cultivated germplasm around the world. These results demonstrate that the markers based on intron, exon and promoter regions of *ALMT1* can trace the inheritance of the Al resistance locus within wheat pedigrees and track Al resistance in breeding programmes.

In barley, Al tolerance is conditioned by the *Alp* locus on chr. arm 4HL, which is associated with Al-activated release of citrate from roots (Raman et al. 2002, 2003; Wang et al. 2006). Among barley genotypes, the cv. Dayton has been reported to exhibit the highest level of Al tolerance (Raman et al. 2003). Using 2,070 F₂ plants from a cross between Dayton and the Al-sensitive cv. Gardiner, Wang et al. (2007) fine-mapped the *Alp* locus to a 0.2 cM interval that includes the *HvMATE* gene. Relative expression of the *HvMATE* gene was

30-fold greater in Dayton than in Gardiner. Additionally, *HvMATE* expression in the F_{2:3} families tested, was significantly correlated with Al tolerance and Al-activated citrate efflux of the F_{2:3} families tested. These results identify *HvMATE*, a gene encoding a multidrug and toxic compound extrusion protein, as a suitable candidate controlling Al tolerance in barley. In both barley and wheat, the mRNA expression levels of the cloned tolerance genes was strongly and positively correlated with tolerance and organic acid secretion levels in collections of diverse genotypes of these species (Raman et al. 2005; Furukawa et al. 2007; Wang et al. 2007).

18.3.5 Tolerance to Boron Toxicity

As an essential micronutrient, boron (B) can curtail Triticeae yield either because it is present in a limiting (Rerkasem et al. 2004) or excessive (Yau et al. 1997; Kalayci et al. 1998) amount in the soil. An excessive concentration of B in soils occurs naturally in semiarid regions in Australia, West Asia and North Africa, and can be caused by irrigation water high in B (Collins et al. 2008). In barley, QTLs for tolerance to high soil B were mapped on chrs. 2H, 3H and 4H of a population of DH lines derived from Clipper \times Sahara 3771 (Jefferies et al. 1999; Karakousis et al. 2003). Recently, the major QTL on chr. 4H has been fine-mapped and cloned (Sutton et al. 2007), revealing that Sahara 3771, the B-tolerant parental line, contains about four times as many *Bot1* gene copies and produces substantially more *Bot1* transcript. Additionally, the *Bot1* protein from Sahara 3771 showed a higher capacity to provide B tolerance in yeast. At the plant level, *Bot1* regulates the net entry of B into the root and its disposal from leaves via guttation. The cloning of *Bot1* provides new opportunities for a more accurate and effective manipulation of B tolerance in barley. In fact, tight linkage between *Bot1* and undesirable alleles from Sahara may have discouraged the utilisation of this B toxicity tolerance through MABC based on the markers flanking the *Bot1* locus. In hexaploid wheat, genetic variation for boron toxicity has been well-documented (Ahmed et al. 2007). The evaluation under B-toxic conditions of 161 DH lines from a cross between the B-tolerant cv. Halberd and the moderately B-sensitive cv. Cranbrook allowed Jefferies et al. (2000) to identify two QTLs on chrs. 7BL and 7D. While both QTLs were associated with leaf symptom expression, the 7BL QTL (*Bo1*) was also associated with the control of B uptake and with a reduction in the effect of boron toxicity on root-growth suppression. Importantly, RFLP markers at both QTLs were shown to be effective in selecting for improved B tolerance in an alternative genetic background, hence paving the way to use MAS for improving B tolerance (Jefferies et al. 2000; Langridge 2005). Subsequently, Schnurbusch et al. (2007) deployed rice-wheat synteny to fine map *Bo1* to a 1.8 cM interval which corresponds to a 227-kb section of rice chr. arm 6L encoding 21 predicted proteins with no homology to any known B transporters.

Interestingly, a co-dominant PCR marker *AWW5L7* co-segregated with *Bo1* and was highly predictive of the level of tolerance to B of a set of 94 Australian bread wheat accessions. This new marker is presently being exploited for breeding purposes to improve B tolerance in Australian bread wheat (T. Schnurbusch and P. Langridge, personal communication).

To further investigate the molecular basis of B phytotoxicity in barley, Patterson et al. (2007) used iTRAQ peptide tagging (iTRAQ) to compare the abundances of proteins from B-tolerant and B-intolerant barley plants from a Clipper × Sahara 3771 doubled-haploid population selected on the basis of a presence or absence of two B-tolerance QTLs. In particular, B-tolerant plants showed an elevated level of three enzymes involved in siderophore production, namely Iron Deficiency Sensitive2 (IDS2), IDS3 and a methylthio-ribose kinase, suggesting a potential link between iron, B and the siderophore hydroxymugineic acid.

A metabolomics approach was deployed by Roessner et al. (2006) in barley to compare metabolite profiles in root and leaf tissues of B-intolerant cv. Clipper and B-tolerant Sahara 3771, an Algerian landrace. Exposure to elevated B (200 and 1,000 mM), caused greater differences between tissue types (e.g. roots and leaves) than between cultivars. As compared to sensitive cultivars, Sahara maintains a significantly lower root B concentration in roots, xylem sap and leaves. Additionally, the ability of Sahara to maintain lower root B concentrations was constitutive and occurred across a wide range of B levels. Interestingly, this ability was totally lost at high pH (Hayes and Reid 2004), an example of how combining different stresses can change the level of tolerance shown to each single stress.

18.3.6 Tolerance to Zinc and Manganese Deficiency

Among the most widespread crop micronutrient deficiencies, zinc (Zn) deficiency afflicts large areas causing severe yield losses. Additionally, Zn deficiency negatively affects human nutrition and health in many developing countries. In wheat, tolerance to Zn-deficient soils correlates with high-Zn content in the grain, suggesting that the two traits can be selected simultaneously (Graham et al. 1999; Gelin et al. 2007). To date, QTLs for Zn efficiency have been described in rice (Wissuwa et al. 2006) but not in the Triticeae. It should be noted that in some wheat genotypes it has been shown that soil moisture may facilitate acquisition of Zn (Bagci et al. 2007), a clear example on how different abiotic stresses can interact with one another.

Wheat chr. 6B derived from wild emmer wheat (*Triticum turgidum* ssp. *dicoccoides*) was shown to be associated with high Zn concentration in the grain. Distelfeld et al. (2007) used recombinant chromosome substitution lines (RSLs), previously constructed for genetic and physical maps of *Gpc-B1* (a 250-kb locus affecting grain protein concentration), to evaluate the effects of

the *Gpc-B1* locus on grain micronutrient concentrations. RSLs carrying the *Gpc-B1* allele of *T. dicoccoides* accumulated on average higher concentrations of Zn (+12%), Fe (+18%), Mn (+29%) and protein in the grain (+38%) compared with RSLs carrying the allele from durum wheat. Additionally, the high grain Zn, Fe and Mn concentrations were consistently expressed across five environments. These findings also confirmed the previously reported effect of the wild-type allele of *Gpc-B1* on earlier senescence of flag leaves (Uauy et al. 2006). Based on these results, the *Gpc-B1* locus appears to regulate remobilisation of protein, zinc, iron and manganese from leaves to the grains.

A peculiar interaction between Zn and phosphate (P) uptake was evidenced in barley (Huang et al. 2000). Because P uptake is regulated through high-affinity P transporter proteins embedded in the plasma membrane of certain cell types in roots, expression of the genes that encode these transporters responds to the P status of the plants, and their transcription is normally tightly controlled. Under Zn deficiency, this tight control of P uptake is lost, thus leading to very high accumulation of P. Accordingly, Huang et al. (2000) showed that in barley the expression of the HVPT1 and HVPT2 high-affinity P transporters was induced by Zn deficiency in plants grown in either P-sufficient or -deficient conditions. Interestingly, the role of Zn in the regulation of these genes was found to be specific in that it could not be replaced by manganese, a divalent cation similar to Zn. These results suggested that Zn plays a specific role in the signal transduction pathway responsible for the regulation of genes encoding high-affinity P transporters in plant roots (Huang et al. 2000).

Among the micronutrient deficiency that curtails barley and wheat production, an important role is also played by manganese (Mn). In barley, Pallotta et al. (2000) used a pot-based bioassay and bulked segregant analysis to identify RFLPs linked to a locus (designated as *Mell*) for Mn efficiency found to be distally located on chr. 4HS. Field evaluation of marker-selected progeny supported the major role of *Mell* in the genetic control of Mn efficiency, which led to the adoption of marker-assisted selection for this trait in the Southern Australian barley breeding program (Pallotta et al. 2000).

18.3.7 Tolerance to Waterlogging

In marginal and high rainfall zones, soil waterlogging is a widespread environmental constraint limiting wheat and barley productivity (Zhang et al. 2006; Kubo et al. 2007). Because genetic differences in tolerance at the germination stage often differ from tolerance at later stages of development, the current view is that different mechanisms of tolerance exist at the whole plant and tissue levels (Setter and Waters 2003). The physiological mechanisms imparting tolerance to waterlogging can be related to adaptive traits relating to phenology, morphology and anatomy, nutrition, metabolism (e.g. anaerobic catabolism) and post-anoxic damage and recovery. After reviewing the existing literature,

Setter and Waters (2003) concluded that the best opportunities for improving germplasm for tolerance to waterlogging are for further exploration and utilisation of genetic diversity by improving selection criteria including the use of MAS. To date, only Burgos et al. (2001) have described QTLs for waterlogging tolerance in the Triticeae by using an RIL population derived from a cross between cv. Forno (*Triticum aestivum*) and cv. Oberkulmer (*Triticum spelta*). Five QTLs accounting for 41% of the phenotypic variance for survival to 48 h of flooding were mapped on chrs. 2B, 3B, 5A and 7D. Plants with a fast coleoptile growth were more tolerant to flooding, as indicated by the fact that flooding tolerance was highly correlated ($r = 0.80$) with germination time. For seedling growth index after flooding, ten QTLs localised on chrs. 2A, 2B, 2D, 3A, 4B, 5A, 5B, 6A and 7D accounted for 36% of the phenotypic variance.

To evaluate the response of wheat roots to a controlled hypoxic environment at the transcription level, Lee et al. (2007) assembled a cDNA library from hypoxia-stressed wheat roots. Starting from a total of 1,274 clones of at least 50-bp long, a BLASTx analysis against the non-redundant database identified 494 unigenes with high homology ($E < e^{-10}$) to the registered plant amino acid sequences. Additionally, 112 unigenes were not matched to any wheat ESTs in the database. These ESTs may provide useful information for further investigations to better understand the response of wheat roots to waterlogging. Potential sources of valuable allelic variability for root features (e.g. formation of aerenchima; Thomson et al. 1990) able to influence tolerance to waterlogging include wild relatives of wheat (van Ginkel and Ogbonnaya 2007) and barley (Garthwaite et al. 2003).

18.3.8 Tolerance to Low Temperature

Among all crops, Triticeae are probably exposed to the widest range and daily fluctuations in temperature. As an example, barley and wheat cultivation spans from the sizzling hot and arid lowlands of Mediterranean countries to the cold regions at high elevations or in plains at high latitude in the Northern hemisphere. Additionally, frost can drastically curtail barley and wheat production in temperate regions and at high elevations in the tropics due to a reduced fertility caused by low temperature at flowering (Fuller et al. 2007). In southern Australia, spring radiation frost reduces yield potential and grain quality of barley by damaging sensitive reproductive organs in the latter stages of development (Reinheimer et al. 2004).

In hexaploid wheat, the early studies of Sutka and co-workers using chromosome substitution lines pointed out the quantitative nature of frost tolerance (Sutka 1994) and provided a useful framework for more informative studies based on QTL analysis. The first QTL study showed that frost tolerance and vernalization requirements in hexaploid wheat are governed by two tightly linked loci on chr. 5A, namely *Vrn-1* and *Fr1* (Galiba et al. 1995; Sutka et al.

1999). An additional frost tolerance locus, *Fr2*, was identified on chr. 5D in the homeologous region corresponding to that harbouring *Fr1* on chr. 5A, thus providing evidence for the orthology of these two QTLs (Snape et al. 1997). Subsequently, an additional locus for frost tolerance designated *Fr-A2* was mapped on chr. 5AL of diploid wheat (*Triticum monococcum*), 30 cM proximal to the major frost tolerance locus *Fr-A1* (Vagujfalvi et al. 2003). Remarkably, plants of the frost-tolerant and frost-susceptible parental lines grown at 15 °C differed in the transcription level of the cold-induced gene *Cor14b*. When the transcript level of this gene was determined in each line of the mapping population and mapped as a QTL, its peak precisely overlapped with the QTL peak for frost survival at the *Fr-A2* locus, thus providing strong circumstantial evidence that frost tolerance was mediated by differential regulation of the expression of the *Cor14b* gene. Previous work in hexaploid wheat showed that *Cor14b* is regulated by two loci on chr. 5A, one in the same chromosome region as the *T. monococcum* *Fr-A2* locus and the other one closely linked to *Fr-A1*. Since C-repeat binding factor (CBF) transcriptional activators in Arabidopsis regulate *Cor* genes and are involved in frost tolerance (Shinozaki and Yamaguchi-Shinozaki 2000), Vagujfalvi et al. (2003) mapped the cold-regulated CBF-like barley gene *Chf3* on the *T. monococcum* map. *Chf3* was mapped on the peak of the *Fr-A2* QTL for frost tolerance, suggesting that the observed differential regulation of *Cor14b* at the *Fr-A2* locus is due to allelic variation at the *Chf3* locus, and that this transcriptional activator gene might be a candidate gene for the *Fr-A2* frost tolerance locus on wheat chr. 5A. Recently, Knox et al. (2008) investigated a cluster of 11 CBF genes that was mapped to the Frost resistance-2 (*Fr-Am2*) locus on chr. 5 of *Triticum monococcum*. The *Fr-Am2* locus was mapped at the peak of two overlapping QTLs, one for frost survival and the other for differential expression of the cold regulated gene *COR14b*.

In barley, although early work identified only one major multilocus cluster of linked QTLs on the chr. arm 5HL in the Dicktoo × Morex population (Hayes et al. 1993), nine QTLs for freezing tolerance were later reported on chrs. 2H, 3H, 5H and 6H in the Arda × Opale background (Tuberosa et al. 1997). The presence of a locus (*Dhn1*) encoding for a dehydrin within the QTL interval on chromosome 5H, led to the initial suggestion that this locus might influence frost tolerance (Hayes et al. 1993; Pan et al. 1994), a hypothesis that was later on dismissed based on additional work. Recently, following the identification of two frost tolerance QTLs on the long arm of chr. 5H (*Fr-H1*, distal; *Fr-H2*, proximal) in a Nure × Tremois mapping population, Francia et al. (2007) assembled a large segregating population of 1,849 F₂ plants for the fine mapping of these QTLs, which allowed to confine *Fr-H2* within a 4.6 cM interval. This region harbours several CBF factors that were suggested as likely candidate genes and are being characterised to determine their role (Francia et al. 2007). While the distal loci (*Fr-H1*) are close to or coincident with the *Vrn-1* loci known to control flowering response to vernalization, CBF

gene clusters at the proximal (*Fr-H2*) loci appear as the most likely candidates (Båga et al. 2007; Francia et al. 2007). A cluster of 11 CBF transcription factors is located at the frost tolerance locus *Fr-A(m)2* in *Triticum monococcum* chr. 5A (Miller et al. 2006). These 11 *TmCBF* genes were confined within five different BAC clones with recombination events, which will be useful in identifying which of the CBF genes is responsible for the differences in frost tolerance between the *T. monococcum* parental lines at the *Fr-A(m)2* locus. These results and recent work by Collins and coworkers at ACPFG (N. Collins, personal communication) indicate that positional cloning of frost tolerance QTLs may soon become a reality.

The above-mentioned results clearly indicate that loci affecting vegetative frost tolerance have been identified at corresponding positions across Triticeae at two regions on the long arms of group 5 chromosomes, confirming the findings of previous reviews on the syntenic relationships of QTLs for freezing tolerance in wheat and barley (Snape et al. 2001; Cattivelli et al. 2002). Additionally, transgenesis and comparative genetic studies have shown conserved synteny and functional features for the *CBF* gene family between the Triticeae cereals, perennial ryegrass (Tamura and Yamada 2007) and also rice (Skinner et al. 2005; Gao et al. 2007; Oh et al. 2007), thus suggesting that members of the CBF gene family regulate the stress responses of a wide range of plant species.

18.3.9 Tolerance to High Temperature

During grain filling, Triticeae are frequently exposed to high temperatures that can disrupt the deposition of reserve products, thus lowering final yield and the quality of the grain. Heat stress often accompanies drought stress and the two interact by reducing fertility and grain quality (Barnabás et al. 2008), probably as a result of the disruption of sugar metabolism and growth regulators. The genetic basis of tolerance to heat stress in Triticeae is poorly understood (Dhanda and Munjal 2006; Senthil-Kumar et al. 2007). In wheat, the evaluation of a RIL population derived from Ventnor (heat-tolerant) × Karl 92 (heat-susceptible) allowed Yang et al. (2002) to identify two QTLs on chrs. 1B and 5A for grain filling duration, a trait thought to be correlated with heat tolerance. The effects of high temperature during grain filling on grain yield, kernel weight and GP content were investigated in a population of 194 RILs (Renan × Recital) tested at six locations in France. A varying number (from one to four) of “stable” QTLs (i.e. detected in at least four of the six locations) were identified. Additionally, factorial regression on genotype × environment interaction (GEI) allowed for the determination of some genetic regions involved in the differential reaction of genotypes to specific climatic factors, such as the number of days with a maximum temperature above 25 °C during grain filling.

18.4 Genomics of Genotype \times Environment Interaction Under Conditions of Abiotic Stress

A common finding of studies reporting results of field experiments carried out across a broad range of environmental conditions is the occurrence of GEI for yield and end-use quality, particularly when multiple abiotic stresses occur (Snape et al. 2007). An adequate understanding of GEI is of key importance for developing varieties with both general and specific adaptation. Additionally, a thorough evaluation of the effects of a particular genotype or target locus should duly consider the influence of management factors such as planting density in the field, an agronomic feature that greatly influences crop performance under water- and nutrient-deprived conditions (Reynolds and Tuberosa 2008). A combined QTL analysis for yield of several wheat DH populations evaluated across different environments and seasons allowed Snape et al. (2007) to identify QTLs showing stable and differential expression across irrigated and non-irrigated conditions. Genes on the homoeologous group 2 chromosomes were associated with flag leaf senescence (i.e. stay-green) variation and showed interactions with the water regime. Variation for stem soluble carbohydrate reserves was associated with the 1RS arm of the 1BL/1RS translocated chromosome, and was positively associated with yield under both irrigated and rainfed conditions, thus contributing to general adaptability. Markers associated with major QTLs for grain yield across a broad range of environmental conditions have been described in both bread wheat (Groos et al. 2003; Quarrie et al. 2005; Kirigwi et al. 2007; Kuchel et al. 2007b,c; Laperche et al. 2007; Snape et al. 2007) and durum wheat (Maccaferri et al. 2008).

An extensive QTL study on GEI in bread wheat investigated the performance of 96 DH lines tested for yield and yield components across a range of 24 site \times treatment \times year combinations including drought, salinity and nutrient stress (Quarrie et al. 2005). In total, 17 clusters of yield QTLs were identified in five or more trials. Three of the yield QTL clusters were largely site-specific, while four clusters were prevalently associated with one or other of the stress treatments. Three of the yield QTL clusters overlapped with the dwarfing gene *Rht-B1* on chr. 4BS and with the vernalization genes *Vrn-A1* on chr. 5AL and *Vrn-D1* on chr. 5DL. The QTL clusters with the strongest effects on yield were mapped to chrs. 7AL and 7BL, and were mainly due to variation in grain number per ear.

Other QTL studies have attempted the dissection of GEI using marker \times environmental covariate interactions. In wheat, a study performed with QTL data available for a population of chr. 3A recombinant inbred substitution lines tested in seven environments along with environmental covariate data showed that precipitation and temperature before anthesis had the greatest influence on agronomic performance traits, and explained a sizeable portion of the total GEI for those traits (Campbell et al. 2003). Individual molecular marker \times environmental covariate interactions explained a large portion of the total QTL \times

environment interactions (QEI) for several agronomic traits. In particular, 76% of the QEI for a major grain yield QTL was explained by the effect of temperature during pre-anthesis growth. Therefore, in this study environmental covariates provided a strong basis for explaining QEI using marker \times environmental covariate interactions. A similar study on wheat allowed Groos et al. (2003) to identify stable QTLs (i.e. detected in at least four of the six locations that were tested) for grain protein-content (chrs. 2A, 3A, 4D and 7D), yield (chr. 7D) and kernel weight (on chrs 2B, 5B and 7A).

A population of 182 DH lines progeny derived from Trident \times Molineux was used to characterise across 17 environments the interaction of previously mapped grain yield QTLs (Kuchel et al. 2007c) with specific environmental covariables (e.g. latitude, rainfall, various temperature-based variables, etc.) used for grain yield assessment (Kuchel et al. 2007c). A putative grain yield *per se* QTL on chr. 1B (*QGyld.agt-1B*) evidenced interactions with the amount of winter and spring rainfall, the number of days with maximum temperature above 30 °C, and the number of days with a minimum temperature below 10 °C. Because this locus showed no cross-over interaction effect and the cv. Molineux allele was consistently higher-yielding in response to all environmental covariables, Kuchel et al. (2007a) suggested *QGyld.agt-1B* as a prime candidate for MAS for improved grain yield and wide adaptation in wheat.

18.5 Prospects of Genomics-Assisted Improvement of Tolerance to Abiotic Stress

Despite the spectacular progress of genomics-assisted approaches of the past decade, their impact on the improvement of abiotic stress tolerance of the Triticeae remains marginal, particularly in view of what has been achieved by conventional breeding. This notwithstanding, it is expected that MAS for tailoring cultivars more resilient to environmental constraints will become progressively more prevalent in the years to come (Kuchel et al. 2005; Langridge 2005; Tuberosa et al. 2007; William et al. 2007). The quest and cloning of agronomically valuable loci will be streamlined and made more cost-effective by (i) the improvement of high-throughput phenotyping (Granier et al. 2006; Sadok et al. 2007) and molecular platforms (e.g. SNP and DArT profiling, Kilian 2005; Akbari et al. 2006; Hearnden et al. 2007; Mantovani et al. 2008), (ii) new mapping approaches (e.g. nested-association mapping; Yu et al. 2008), (iii) the increased availability of genomic libraries (Dolezel et al. 2007) and sequence information (e.g. “deep” sequencing of DNA and mRNA; Stein 2007; Pop and Salzberg 2008; Torres et al. 2008), (iv) comparative mapping (Sorrells et al. 2003; Salse and Feuillet 2007) and bioinformatics (Lim et al. 2007; Liang et al. 2008), and (v) genome-wide validation approaches (e.g. RNA interference, Lukens and Zhan 2007; TILLING, Till et al. 2007). Additionally, the cost of MAS (Kuchel et al. 2005, 2007a) will continue to decrease as more

user-friendly marker platforms will allow for increased multiplexing and gel-free profiling (Hayden et al. 2008).

The new paradigm ushered in by the genomics era allows the breeder to focus more effectively on a single locus in order to more exhaustively explore and accurately evaluate the relevant allelic variation (Sørensen et al. 2007). To overcome the shortcomings due to the limited allelic diversity present at key loci, growing attention will be devoted to the screening and characterisation of wild relatives and landraces of the Triticeae crops in an attempt to single out superior alleles that the domestication bottleneck and modern agriculture have left behind (Reynolds et al. 2007b; Feuillet et al. 2008). Advanced-backcross QTL analysis (Tanksley and Nelson 1996; Grandillo et al. 2007) and introgression libraries will be instrumental for identifying agronomically superior alleles absent within the elite germplasm. Modeling based on a simplified framework of the components interacting at the physiological and developmental level is expected to streamline the identification of the alleles that should preferentially be pyramided through MAS to improve stress-adaptation and yield stability (Tardieu 2003; Yin et al. 2003; van Eeuwijk et al. 2005; Hammer et al. 2006; Cooper et al. 2007; Welcker et al. 2007). Phenotyping large collections of unrelated accessions across multiple environments (i.e. combinations of locations, years and management practices) will allow modeling GEI (Crossa et al. 2007) and better classification of the nature, adaptive or constitutive, of QTLs.

Increasing evidence suggests a prominent role of epigenesis (Lukens and Zhan 2007; Pfluger and Wagner 2007; Boyko and Kovalchuk 2008) and RNA interference (Chinnusamy et al. 2007; Jung et al. 2007; Sunkar et al. 2007) in the regulation of gene expression under stress conditions. The dissection of functional differences between QTL alleles, including variation in non-coding regions (Magalhaes et al. 2007; Salvi et al. 2007) will allow for a more complete understanding of natural variation and its role in sustaining crop performance in harsh conditions. Integration of QTL information into a breeding pipeline aimed at improving tolerance to abiotic stresses will best be achieved within a multidisciplinary context able to link the stress-responsive mechanisms of crops with the functional variation of the relevant networks at the cellular and molecular level (Wollenweber et al. 2005). Nonetheless, if the effects of the loci affecting the target trait are too small, it will be more cost-effective to simply select phenotypically for yield itself, rather than attempting MAS. In terms of target traits for MAS and other genomics approaches, root architectural traits appear ideal targets because of the difficulty to visually select for root features (de Dorlodot et al. 2007) and the pivotal role that root architecture plays in almost all abiotic stresses, as shown by the results surveyed in this review.

The results presented in this review indicate that genomics approaches now allow for the fine mapping and cloning of QTLs for tolerance to abiotic stress in the Triticeae (Vagujfalvi et al. 2003; Huang et al. 2006; Schnurbusch et al. 2007; Sutton et al. 2007), opening up new opportunities for a more accurate manipulation of the genes governing the adaptive response to environmental

constraints. Ultimately, the practical impact of the knowledge and materials acquired through genomics-based approaches will depend on their integration and exploitation within the available phenotypic variation and practices of breeding programs.

References

- Ahmed, M., Jahiruddin, M. and Mian, M.H. (2007) Screening of wheat genotypes for boron efficiency. *J. Plant Nutr.* 30, 1127–1138.
- Akbari, M., Wenzl, P., Caig, V., Carling, J., Xia, L., Yang, S.Y., et al. (2006) Diversity arrays technology (DArT) for high-throughput profiling of the hexaploid wheat genome. *Theor. Appl. Genet.* 113, 1409–1420.
- Alonso, J.M. and Ecker, J.R. (2006) Moving forward in reverse: genetic technologies to enable genome-wide phenomic screens in *Arabidopsis*. *Nat. Rev. Genet.* 7, 524–536.
- An, D.G., Su, J.Y., Liu, Q.Y., Zhu, Y.G., Tong, Y.P., Li, J.M., et al. (2006) Mapping QTLs for nitrogen uptake in relation to the early growth of wheat (*Triticum aestivum* L.). *Plant Soil* 284, 73–84.
- Andersen, J.R. and Lubberstedt, T. (2003) Functional markers in plants. *Trends Plant Sci.* 8, 554–560.
- Aniol, A. and Gustafson, J.P. (1984) Chromosome location of genes controlling aluminum tolerance in wheat, rye and triticale. *Can. J. Genet. Cytol.* 26, 701–705.
- Aniol, A. (2004) Chromosomal location of aluminium tolerance genes in rye. *Plant Breed.* 123, 132–136.
- Araus, J.L., Slafer, G.A., Reynolds, M.P. and Royo, C. (2002) Plant breeding and drought in C-3 cereals: what should we breed for? *Ann. Bot.* 89, 925–940.
- Ashraf, M. (2002) Salt tolerance of cotton: some new advances. *Crit. Rev. Plant Sci.* 21, 1–30.
- Atienza, S.G., Faccioli, P., Perrotta, G., Dalfino, G., Zschiesche, W., Humbeck, K., et al. (2004) Large scale analysis of transcripts abundance in barley subjected to several single and combined abiotic stress conditions. *Plant Sci.* 167, 1359–1365.
- Badawi, M., Danyluk, J., Boucho, B., Houde, M. and Sarhan, F. (2007) The CBF gene family in hexaploid wheat and its relationship to the phylogenetic complexity of cereal CBFs. *Mol. Genet. Genomics* 277, 533–554.
- Băga, M., Chodaparambil, S.V., Limin, A.E., Pecar, M., Fowler, D.B. and Chibbar, R.N. (2007) Identification of quantitative trait loci and associated candidate genes for low-temperature tolerance in cold-hardy winter wheat. *Funct. Integr. Genomics* 7, 53–68.
- Bagci, S.A., Ekiz, H., Yilmaz, A. and Cakmak, I. (2007) Effects of zinc deficiency and drought on grain yield of field-grown wheat cultivars in Central Anatolia. *J. Agron. Crop Sci.* 193, 198–206.
- Bage, M., Xia, X.C. and Lubberstedt, T. (2007) Functional markers in wheat – commentary. *Curr. Opin. Plant Biol.* 10, 211–216.
- Bahieldin, A., Mahfouz, H.T., Eissa, H.F., Saleh, O.M., Ramadan, A.M., Ahmed, I.A., et al. (2005) Field evaluation of transgenic wheat plants stably expressing the HVA1 gene for drought tolerance. *Physiol. Plantarum* 123, 421–427.
- Bahrman, N., Le Gouis, J., Negroni, L., Amilhat, L., Leroy, P., Laine, A.L., et al. (2004) Differential protein expression assessed by two-dimensional gel electrophoresis for two wheat varieties grown at four nitrogen levels. *Proteomics* 4, 709–719.
- Bahrman, N., Gouy, A., Devienne-Barret, F., Hirel, B., Vedele, F. and Le Gouis, J. (2005) Differential change in root protein patterns of two wheat varieties under high and low nitrogen nutrition levels. *Plant Sci.* 168, 81–87.
- Barnabás, B., Jager, K. and Feher, A. (2008) The effect of drought and heat stress on reproductive processes in cereals. *Plant Cell Environ.* 31, 11–38.

- Baum, M., Grando, S., Backes, G., Jahoor, A., Sabbagh, A. and Ceccarelli, S. (2003) QTLs for agronomic traits in the Mediterranean environment identified in recombinant inbred lines of the cross 'Arta' \times *H. spontaneum* 41-1. *Theor. Appl. Genet.* 107, 1215–1225.
- Bernard, S.M., Møller, A.L.B., Dionisio, G., Kichey, T., Jahn, T.P., Dubois, F., et al. (2008) Gene expression, cellular localisation and function of glutamine synthetase isozymes in wheat (*Triticum aestivum* L.). *Plant Mol. Biol.* DOI 10.1007/s11103-008-9303-y.
- Blum, A. (1988) *Breeding for Stress Environments*. CRC Press, Boca Raton.
- Blum, A. (1996) Crop responses to drought and the interpretation of adaptation. *Plant Growth Regul.* 20, 135–148.
- Blum, A. (2005) Drought resistance, water-use efficiency, and yield potential – are they compatible, dissonant, or mutually exclusive? *Aust. J. Agric. Res.* 56, 1159–1168.
- Bohnert, H.J., Gong, Q.Q., Li, P.H. and Ma, S.S. (2006) Unraveling abiotic stress tolerance mechanisms – getting genomics going. *Curr. Opin. Plant Biol.* 9, 180–188.
- Boisson, M., Mondon, K., Torney, V., Nicot, N., Laine, A.L., Bahrman, N., et al. (2005) Partial sequences of nitrogen metabolism genes in hexaploid wheat. *Theor. Appl. Genet.* 110, 932–940.
- Borlaug, N.E. and Dowsell, C.R. (2005) Feeding a world of ten billion people: a 21st century challenge. In: R. Tuberosa, R.L. Phillips and M. Gale (Eds.), *Proceedings of the International Congress "In the Wake of the Double Helix: From the Green Revolution to the Gene Revolution"*, 27–31 May 2003, Bologna, Italy. Avenue media, Bologna, pp. 3–24.
- Borlaug, N.E. (2007) Sixty-two years of fighting hunger: personal recollections. *Euphytica* 157, 287–297.
- Boyko, A. and Kovalchuk, I. (2008) Epigenetic control of plant stress response. *Environ. Mol. Mutagen.* 49, 61–72.
- Breseghele, F. and Sorrells, M.E. (2006) Association analysis as a strategy for improvement of quantitative traits in plants. *Crop Sci.* 46, 1323–1330.
- Brini, F., Hanin, M., Mezghani, I., Berkowitz, G.A. and Masmoudi, K. (2007) Overexpression of wheat Na^+/H^+ antiporter *TNAX1* and H^+ -pyrophosphatase *TVPI* improve salt- and drought-stress tolerance in *Arabidopsis thaliana* plants. *J. Exp. Bot.* 58, 301–308.
- Buckler, E.S.I. and Thornsberry, J.M. (2002) Plant molecular diversity and applications to genomics. *Curr. Opin. Plant Biol.* 5, 107–111.
- Burgos, M.S., Messmer, M.M., Stamp, P. and Schmid, J.E. (2001) Flooding tolerance of spelt (*Triticum spelta* L.) compared to wheat (*Triticum aestivum* L.) – a physiological and genetic approach. *Euphytica* 122, 287–295.
- Burke, J.M., Burger, J.C. and Chapman, M.A. (2007) Crop evolution: from genetics to genomics. *Curr. Opin. Genet. Dev.* 17, 525–532.
- Byrt, C.S., Platten, J.D., Spielmeier, W., James, R.A., Lagudah, E.S., Dennis, E.S., et al. (2007) HKT1;5-like cation transporters linked to Na^+ exclusion loci in wheat, *Nax2* and *Knal*. *Plant Physiol.* 143, 1918–1928.
- Caldwell, D.G., McCallum, N., Shaw, P., Muehlbauer, G.J., Marshall, D.F. and Waugh, R. (2004) A structured mutant population for forward and reverse genetics in Barley (*Hordeum vulgare* L.). *Plant J.* 40, 143–150.
- Campbell, B.T., Baenziger, P.S., Eskridge, K.M., Budak, H., Streck, N.A., Weiss, A., et al. (2003) Using environmental covariates to explain genotype \times environment and QTL \times environment interactions for agronomic traits on chromosome 3A of wheat. *Crop Sci.* 44, 620–627.
- Campbell, S.A. and Close, T.J. (1997) Dehydrins: genes, proteins, and associations with phenotypic traits. *New Phytol.* 137, 61–74.
- Cattivelli, L., Baldi, P., Crosatti, C., Di Fonzo, N., Faccioli, P., Grossi, M., et al. (2002a) Chromosome regions and stress-related sequences involved in resistance to abiotic stress in Triticeae. *Plant Mol. Biol.* 48, 649–665.
- Cattivelli, L., Rizza, F., Badeck, F.W., Mazzucotelli, E., Mastrangelo, A.M., Francia, E., et al. (2008) Drought tolerance improvement in crop plants: an integrated view from breeding to genomics. *Field Crops Res.* 105, 1–14.

- Ceccarelli, S. and Grando, S. (1996) Drought as a challenge for the plant breeder. *Plant Growth Regul.* 20, 149–155.
- Chauvin, L.P., Houde, M. and Sarhan, F. (1993) A leaf-specific gene stimulated by light during wheat acclimation to low temperature. *Plant Mol. Biol.* 23, 255–265.
- Chen, Z., Newman, I., Zhou, M., Mendham, N., Zhang, G. and Shabala, S. (2005) Screening plants for salt tolerance by measuring K^+ flux: a case study for barley. *Plant Cell Environ.* 28, 1230–1246.
- Chen, Z.H., Pottosin, I.I., Cuin, T.A., Fuglsang, A.T., Tester, M., Jha, D., et al. (2007a) Root plasma membrane transporters controlling K^+/Na^+ homeostasis in salt-stressed barley. *Plant Physiol.* 145, 1714–1725.
- Chen, Z.H., Zhou, M.X., Newman, I.A., Mendham, N.J., Zhang, G.P. and Shabala, S. (2007b) Potassium and sodium relations in salinised barley tissues as a basis of differential salt tolerance. *Funct. Plant Biol.* 34, 150–162.
- Chinnusamy, V., Zhu, J., Zhou, T. and Zhu, J.-K. (2007) Small RNAs: big role in abiotic stress tolerance of plants. In: M.A. Jenks, P.M. Hasegawa and S. Mohan Jain (Eds.), *Advances in Molecular Breeding Toward Drought and Salt Tolerant Crops*. Springer, Dordrecht, pp. 223–260.
- Choi, D.W., Zhu, B. and Close, T.J. (1999) The barley (*Hordeum vulgare* L.) dehydrin multigene family: sequences, allele types, chromosome assignments, and expression characteristics of 11 *Dhn* genes of cv Dicktoo. *Theor. Appl. Genet.* 98, 1234–1247.
- Choi, D.W., Koag, M.C. and Close, T.J. (2000) Map locations of barley *Dhn* genes determined by gene-specific PCR. *Theor. Appl. Genet.* 101, 350–354.
- Choi, D.W., Rodriguez, E.M. and Close, T.J. (2002) Barley *Cbf3* gene identification, expression pattern, and map location. *Plant Physiol.* 129, 1781–1787.
- Collins, N., Tardieu, F. and Tuberosa, R. (2008) QTL approaches for improving crop performance under abiotic stress conditions: where do we stand? *Plant Physiol.* 147, 469–486.
- Colmer, T.D., Flowers, T.J. and Munns, R. (2006) Use of wild relatives to improve salt tolerance in wheat. *J. Exp. Bot.* 57, 1059–1078.
- Comai, L., Young, K., Till, B.J., Reynolds, S.H., Greene, E.A., Codomo, C.A., et al. (2004) Efficient discovery of DNA polymorphisms in natural populations by Ecotilling. *Plant J.* 37, 778–786.
- Condon, A.G., Richards, R.A., Rebetzke, G.J. and Farquhar, G.D. (2004) Breeding for high water-use efficiency. *J. Exp. Bot.* 55, 2447–2460.
- Consoli, L., Lefevre, A., Zivy, M., de Vienne, D. and Damerval, C. (2002) QTL analysis of proteome and transcriptome variations for dissecting the genetic architecture of complex traits in maize. *Plant Mol. Biol.* 48, 575–581.
- Cooper, M., Podlich, D.W. and Luo, L. (2007) Modeling QTL effects and MAS in plant breeding. In: R.K. Varshney and R. Tuberosa (Eds.), *Genomics-Assisted Crop Improvement-Vol 1: Genomics Approaches and Platforms*. Springer, Dordrecht, pp. 57–95.
- Crosatti, C., Mare, C., Mazzucotelli, E., Selioni, S., Barilli, S., Bassi, R., et al. (2003) Genetic analysis of the expression of the cold-regulated gene *cor14b*: a way toward the identification of components of the cold response signal transduction in Triticeae. *Can. J. Bot.* 81, 1162–1167.
- Crosatti, C., Pagani, D., Cattivelli, L., Stanca, A.M. and Rizza, F. (2008) Effects of growth stage and hardening conditions on the association between frost resistance and the expression of the cold-induced protein COR14b in barley. *Environ. Exp. Bot.* 62, 93–100.
- Crossa, J., Burgueno, J., Dreisigacker, S., Vargas, M., Herrera-Foessel, S.A., Lillemo, M., et al. (2007) Association analysis of historical bread wheat germplasm using additive genetic covariance of relatives and population structure. *Genetics* 177, 1889–1913.
- Damerval, C., Maurice, A., Josse, J.M. and de Vienne, D. (1994) Quantitative trait loci underlying gene product variation: a novel perspective for analyzing regulation of genome expression. *Genetics* 137, 289–301.

- Davenport, R., James, R.A., Zakrisson-Plogander, A., Tester, M. and Munns, R. (2005) Control of sodium transport in durum wheat. *Plant Physiol.* 137, 807–818.
- de Dorlodot, S., Forster, B., Pages, L., Price, A., Tuberosa, R. and Draye, X. (2007) Root system architecture: opportunities and constraints for genetic improvement of crops. *Trends Plant Sci.* 12, 474–481.
- Delhaize, E., Ryan, P.R., Hebb, D.M., Yamamoto, Y., Sasaki, T. and Matsumoto H. (2004) Engineering high-level aluminum tolerance in barley with the *ALMT1* gene. *Proc. Natl. Acad. Sci. USA* 101, 15249–15254.
- Delhaize, E., Gruber, B.D. and Ryan, P.R. (2007) The roles of organic anion permeases in aluminium resistance and mineral nutrition. *FEBS Lett.* 581, 2255–2262.
- Desnos, T. (2008) Root branching responses to phosphate and nitrate. *Curr. Opin. Plant Biol.* 11, 82–87.
- Dhanda, S.S. and Munjal, R. (2006) Inheritance of cellular thermotolerance in bread wheat. *Plant Breed.* 125, 557–564.
- Diab, A.A., Teulat-Merah, B., This, D., Ozturk, N.Z., Benscher, D. and Sorrells, M.E. (2004) Identification of drought-inducible genes and differentially expressed sequence tags in barley. *Theor. Appl. Genet.* 109, 1417–1425.
- Distelfeld, A., Cakmak, I., Peleg, Z., Ozturk, L., Yazici, A.M., Budak, H., et al. (2007) Multiple QTL-effects of wheat *Gpc-B1* locus on grain protein and micronutrient concentrations. *Physiol. Plantarum* 129, 635–643.
- Doerge, R.W. (2002) Mapping and analysis of quantitative trait loci in experimental populations. *Nat. Rev. Genet.* 3, 43–52.
- Doerge, R.W. (2007) Statistical advances in functional genomics. In: R.K. Varshney and R. Tuberosa (Eds.), *Genomics-Assisted Crop Improvement-Vol 1: Genomics Approaches and Platforms*. Springer, Dordrecht, pp. 315–332.
- Dolezel, J., Kubalakova, M., Paux, E., Bartos, J. and Feuillet, C. (2007) Chromosome-based genomics in the cereals. *Chromosome Res.* 15, 51–66.
- Dong, G.Q., Ni, Z.F., Yao, Y.Y., Nie, X.L. and Sun, Q.X. (2007) Wheat Dof transcription factor WPBF interacts with *TaQM* and activates transcription of an alpha-gliadin gene during wheat seed development. *Plant Mol. Biol.* 63, 73–84.
- Dubcovsky, J., Maria, G.S., Epstein, E., Luo, M.C. and Dvorak, J. (1996) Mapping of the K^+/Na^+ discrimination locus *Kna1* in wheat. *Theor. Appl. Genet.* 92, 448–454.
- Ehdaie, B., Whitkus, R.W. and Waines, J.G. (2003) Root biomass, water-use efficiency, and performance of wheat-rye translocations of chromosomes 1 and 2 in spring bread wheat 'Pavon'. *Crop Sci.* 43, 710–717.
- Ehdaie, B., Alloush, G.A. and Waines, J.G. (2008) Genotypic variation in linear rate of grain growth and contribution of stem reserves to grain yield in wheat. *Field Crops Res.* 106, 34–43.
- Ellis, R.P., Forster, B.P., Gordon, D.C., Handley, L.L., Keith, R.P., Lawrence, P., et al. (2002) Phenotype/genotype associations for yield and salt tolerance in a mapping population segregating for two dwarfing genes. *J. Exp. Bot.* 53, 1–14.
- Ershov, P.V., Vasekina, A.V., Voblikova, V.D., Taranov, V.V., Roslyakova, T.V. and Babakov, A.V. (2007) Identification of K^+/H^+ antiporter homolog in barley: expression in cultivars with different tolerance to NaCl. *Russ. J. Plant Physiol.* 54, 16–24.
- Ersoz, E.S., Yu, J. and Buckler, E.S. (2007) Applications of linkage disequilibrium and association mapping. In: R.K. Varshney and R. Tuberosa (Eds.), *Genomics-Assisted Crop Improvement-Vol 1: Genomics Approaches and Platforms*. Springer, Dordrecht, pp. 97–120.
- Faccioli, P., Lagonigro, M.S., De Cecco, L., Stanca, A.M., Alberici, R. and Terzi, V. (2002) Analysis of differential expression of barley ESTs during cold acclimatization using microarray technology. *Plant Biol.* 4, 630–639.
- Fan, X.R., Gordon-Weeks, R., Shen, Q.R. and Miller, A.J. (2006) Glutamine transport and feedback regulation of nitrate reductase activity in barley roots leads to changes in cytosolic nitrate pools. *J. Exp. Bot.* 57, 1333–1340.

- Fan, X.W., Li, F.M., Xiong, Y.C., An, L.Z. and Long, R.J. (2008) The cooperative relation between non-hydraulic root signals and osmotic adjustment under water stress improves grain formation for spring wheat varieties. *Physiol. Plantarum* 132, 283–292.
- Farajalla, M.R. and Gulick, P.J. (2007) The alpha-tubulin gene family in wheat (*Triticum aestivum* L.) and differential gene expression during cold acclimation. *Genome* 50, 502–510.
- Federico, M.L., Kaeppler, H.F. and Skadsen, R.W. (2005) The complex developmental expression of a novel stress-responsive barley *Ltp* gene is determined by a shortened promoter sequence. *Plant Mol. Biol.* 57, 35–51.
- Feuillet, C., Langridge, P. and Waugh, R. (2008) Cereal breeding takes a walk on the wild side. *Trends Genet.* 24, 24–32.
- Finkemeier, I., Kluge, C., Metwally, A., Georgi, M., Grotjohann, N. and Dietz, K.J. (2003) Alterations in Cd-induced gene expression under nitrogen deficiency in *Hordeum vulgare*. *Plant Cell Environ.* 26, 821–833.
- Flint-Garcia, S.A., Thornsberry, J.M. and Buckler, E.S. (2003) Structure of linkage disequilibrium in plants. *Annu. Rev. Plant Biol.* 54, 357–374.
- Flowers, T.J. and Yeo, A.R. (1995) Breeding for salinity tolerance in crop plants: where next? *Aust. J. Plant Physiol.* 22, 875–884.
- Fontecha, G., Silva-Navas, J., Benito, C., Mestres, M.A., Espino, F.J., Hernandez-Riquer, M.V., et al. (2007) Candidate gene identification of an aluminum-activated organic acid transporter gene at the *Alt4* locus for aluminum tolerance in rye (*Secale cereale* L.). *Theor. Appl. Genet.* 114, 249–260.
- Forster, B.P., Ellis, R.P., Thomas, W.T.B., Newton, A.C., Tuberosa, R., This, D., et al. (2000) The development and application of molecular markers for abiotic stress tolerance in barley. *J. Exp. Bot.* 51, 19–27.
- Forster, B.P., Ellis, R.P., Moir, J., Talamè, V., Sanguineti, M.C., Tuberosa, R., et al. (2004) Genotype and phenotype associations with drought tolerance in barley tested in North Africa. *Ann. Appl. Biol.* 144, 157–168.
- Francia, E., Rizza, F., Cattivelli, L., Stanca, A.M., Galiba, G., Toth, B., et al. (2004) Two loci on chromosome 5H determine low-temperature tolerance in a ‘Nure’ (winter) × ‘Tremois’ (spring) barley map. *Theor. Appl. Genet.* 108, 670–680.
- Francia, E., Barabaschi, D., Tondelli, A., Laido, G., Rizza, F., Stanca, A.M., et al. (2007) Fine mapping of a HvCBF gene cluster at the frost resistance locus *Fr-H2* in barley. *Theor. Appl. Genet.* 115, 1083–1091.
- Fresneau, C., Ghashghaie, J. and Cornic, G. (2007) Drought effect on nitrate reductase and sucrose-phosphate synthase activities in wheat (*Triticum durum* L.): role of leaf internal CO₂. *J. Exp. Bot.* 58, 2983–2992.
- Fu, D.L., Uauy, C., Blechl, A. and Dubcovsky, J. (2007) RNA interference for wheat functional gene analysis. *Transgenic Res.* 16, 689–701.
- Fukuda, A., Chiba, K., Maeda, M., Nakamura, A., Maeshima, M. and Tanaka, Y. (2004) Effect of salt and osmotic stresses on the expression of genes for the vacuolar H⁺-pyrophosphatase, H⁺-ATPase subunit A, and Na⁺/H⁺ antiporter from barley. *J. Exp. Bot.* 55, 585–594.
- Fukuda, A. and Tanaka, Y. (2006) Effects of ABA, auxin, and gibberellin on the expression of genes for vacuolar H⁺-inorganic pyrophosphatase, H⁺-ATPase subunit A, and Na⁺/H⁺ antiporter in barley. *Plant Physiol. Biochem.* 44, 351–358.
- Fuller, M.P., Fuller, A.M., Kaniouras, S., Christophers, J. and Fredericks, T. (2007) The freezing characteristics of wheat at ear emergence. *Eur. J. Agron.* 26, 435–441.
- Furukawa, J., Yamaji, N., Wang, H., Mitani, N., Murata, Y., Sato, K., et al. (2007) An aluminum-activated citrate transporter in barley. *Plant Cell Physiol.* 48, 1081–1091.
- Galiba, G., Quarrie, S.A., Sutka, J., Morgounov, A. and Snape, J.W. (1995) RFLP mapping of the vernalization (*Vrn1*) and frost resistance (*Fr1*) genes on chromosome 5A of wheat. *Theor. Appl. Genet.* 90, 1174–1179.

- Galiba, G., Kerepesi, I., Snape, J.W. and Sutka, J. (1997) Location of a gene regulating cold-induced carbohydrate production on chromosome 5A of wheat. *Theor. Appl. Genet.* 95, 265–270.
- Galiba, G., Pecchioni, N., Vágújfalvi, A., Francia, E., Tóth, B., Barabaschi, D., Barilli, S., Crosatti, C., Cattivelli, L. and Stanca, M.A. (2005) Localization of QTLs and candidate genes involved in the regulation of frost resistance in cereals. In: R. Tuberosa, R.L. Phillips and M. Gale (Eds.), *Proceedings of the International Congress "In the Wake of the Double Helix: From the Green Revolution to the Gene Revolution"*, 27–31 May 2003, Bologna, Italy. Avenue media, Bologna, pp. 253–266.
- Gao, J.P., Chao, D.Y. and Lin, H.X. (2007) Understanding abiotic stress tolerance mechanisms: recent studies on stress response in rice. *J. Integr. Plant Biol.* 49, 742–750.
- Garthwaite, A.J., von Bothmer, R. and Colmer, T.D. (2003) Diversity in root aeration traits associated with waterlogging tolerance in the genus *Hordeum*. *Funct. Plant Biol.* 30, 875–889.
- Gelin, J.R., Elias, E.M., Manthey, F.A. and Grant, L. (2007) Study of the relationship between sprouting score and sprout damage in durum wheat (*Triticum turgidum* L. var. *durum*). *Cereal Res. Commun.* 35, 53–61.
- Genc, Y., Huang, C.Y. and Langridge, P. (2007) A study of the role of root morphological traits in growth of barley in zinc-deficient soil. *J. Exp. Bot.* 58, 2775–2784.
- Goff, S.A., Ricke, D., Lan, T.H., Presting, G., Wang, R., Dunn, M., Glazebrook, J., Sessions, A., Oeller, P., Varma, H., et al. (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). *Science* 296, 92–100.
- Gorham, J., Bridges, J., Dubcovsky, J., Dvorak, J., Hollington, P.A., Luo, M.C., et al. (1997) Genetic analysis and physiology of a trait for enhanced K^+/Na^+ discrimination in wheat. *New Phytol.* 137, 109–116.
- Graham, R., Senadhirra, D., Beebe, S., Iglesias, C. and Monasterio, I. (1999) Breeding for micronutrient density in edible portions of staple food crops: conventional approaches. *Field Crops Res.* 60, 57–80.
- Grandillo, S., Tanksley, S.D. and Zamir, D. (2007) Exploitation of natural biodiversity through genomics. In: R.K. Varshney and R. Tuberosa (Eds.), *Genomics-Assisted Crop Improvement-Vol I: Genomics Approaches and Platforms*. Springer, Dordrecht, pp. 121–150.
- Granier, C., Aguirrezabal, L., Chenu, K., Cookson, S.J., Dauzat, M., Hamard, P., et al. (2006) PHENOPSIS, an automated platform for reproducible phenotyping of plant responses to soil water deficit in *Arabidopsis thaliana* permitted the identification of an accession with low sensitivity to soil water deficit. *New Phytol.* 169, 623–635.
- Groos, C., Robert, N., Bervas, E. and Charmet, G. (2003) Genetic analysis of grain protein-content, grain yield and thousand-kernel weight in bread wheat. *Theor. Appl. Genet.* 106, 1032–1040.
- Gulick, P.J., Drouin, S., Yu, Z.H., Danyluk, J., Poisson, G., Monroy, A.F., et al. (2005) Transcriptome comparison of winter and spring wheat responding to low temperature. *Genome* 48, 913–923.
- Gulli, M., Corradi, M., Rampino, P., Marmioli, N. and Perrotta, C. (2007) Four members of the *HSP101* gene family are differently regulated in *Triticum durum* Desf. *Febs Lett.* 581, 4841–4849.
- Guo, P.G., Bai, G.H., Carver, B., Li, R.H., Bernardo, A. and Baum, M. (2007) Transcriptional analysis between two wheat near-isogenic lines contrasting in aluminum tolerance under aluminum stress. *Mol. Genet. Genomics* 277, 1–12.
- Guo, P.G., Baum, M., Varshney, R.K., Graner, A., Grando, S. and Ceccarelli, S. (2008) QTLs for chlorophyll and chlorophyll fluorescence parameters in barley under post-flowering drought. *Euphytica* 163, 203–214.
- Gupta, P.K., Rustgi, S. and Kulwal, P. (2005) Linkage disequilibrium and association studies in higher plants: present status and future prospects. *Plant Mol. Biol.* 57, 461–485.

- Habash, D.Z., Bernard, S., Schondelmaier, J., Weyen, J. and Quarrie, S.A. (2007) The genetics of nitrogen use in hexaploid wheat: N utilisation, development and yield. *Theor. Appl. Genet.* 114, 403–419.
- Hajheidari, M., Eivazi, A., Buchanan, B.B., Wong, J.H., Majidi, I. and Salekdeh, G.H. (2007) Proteomics uncovers a role for redox in drought tolerance in wheat. *J. Proteome Res.* 6, 1451–1460.
- Hammer, G., Cooper, M., Tardieu, F., Welch, S., Walsh, B., van Eeuwijk, F., et al. (2006) Models for navigating biological complexity in breeding improved crop plants. *Trends Plant Sci.* 11, 587–593.
- Haro, R., Banuelos, M.A., Senn, M.A.E., Barrero-Gil, J. and Rodriguez-Navarro, A. (2005) HKT1 mediates sodium uniport in roots. Pitfalls in the expression of HKT1 in yeast. *Plant Physiol.* 139, 1495–1506.
- Hasegawa, P.M., Bressan, R.A., Zhu, J.K. and Bohnert, H.J. (2000) Plant cellular and molecular responses to high salinity. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 51, 463–499.
- Hayden, M.J., Nguyen, T.M., Waterman, A., McMichael, G.L. and Chalmers, K.J. (2008) Application of multiplex-ready PCR for fluorescence-based SSR genotyping in barley and wheat. *Mol. Breed.* 21, 271–281.
- Hayes, J.E. and Reid, R.J. (2004) Boron tolerance in barley is mediated by efflux of boron from the roots. *Plant Physiol.* 136, 3376–3382.
- Hayes, P.M., Blake, T., Chen, T.H.H., Tragoonrung, S., Chen, F., Pan, A., et al. (1993) Quantitative trait loci on barley (*Hordeum vulgare* L.) chromosome 7 associated with components of winterhardiness. *Genome* 36, 66–71.
- Hayes, P.M. and Iyambo, O.E. (1994) Summary of QTL effects in the Steptoe \times Morex population. *Barley Genet. Newsl.* 23, 98–143.
- Hazen, S.P. and Kay, S.A. (2003) Gene arrays are not just for measuring gene expression. *Trends Plant Sci.* 8, 413–416.
- Hazen, S.P., Wu, Y.J. and Kreps, J.A. (2003) Gene expression profiling of plant responses to abiotic stress. *Funct. Integr. Genomics* 3, 105–111.
- Hearnden, P.R., Eckermann, P.J., McMichael, G.L., Hayden, M.J., Eglinton, J.K. and Chalmers, K.J. (2007) A genetic map of 1,000 SSR and DArT markers in a wide barley cross. *Theor. Appl. Gen.* 115, 383–391.
- Hede, A.R., Skovmand, B. and López-Cesati, J. (2001) Acid soil and aluminium toxicity. In: M.P. Reynolds, J.I. Ortiz-Monasterio and A. McNab (Eds.), *Application of Physiology in Wheat Breeding*. D.F. CIMMYT, Mexico, pp. 172–182.
- Hirai, M.Y., Klein, M., Fujikawa, Y., Yano, M., Goodenowe, D.B., Yamazaki, Y., et al. (2005) Elucidation of gene-to-gene and metabolite-to-gene networks in Arabidopsis by integration of metabolomics and transcriptomics. *J. Biol. Chem.* 280, 25590–25595.
- Hong, B.M., Barg, R. and Ho, T.H.D. (1992) Developmental and organ-specific expression of an ABA- and stress-induced protein in barley. *Plant Mol. Biol.* 18, 663–674.
- Houde, M., Danyluk, J., Laliberte, J.F., Rassart, E., Dhindsa, R.S. and Sarhan, F. (1992) Cloning, characterization, and expression of a carrier DNA encoding a 50-kilodalton protein specially induced by cold acclimation in wheat. *Plant Physiol.* 99, 1381–1387.
- Huang, C., Barker, S.J., Langridge, P., Smith, F.W. and Graham, R.G. (2000) Zinc deficiency up-regulates expression of high-affinity phosphate transporter genes in both phosphate-sufficient and -deficient barley roots. *Plant Physiol.* 124, 415–422.
- Huang, C.Y., Roessner, U., Eickmeier, I., Genc, Y., Callahan, D.L., Shirley, N., et al. (2008) Metabolite profiling reveals distinct changes in carbon and nitrogen metabolism in phosphate deficient barley plants (*Hordeum vulgare* L.) *Plant Cell Physiol.* 49, 691–703.
- Huang, S.B., Spielmeier, W., Lagudah, E.S., James, R.A., Platten, J.D., Dennis, E.S., et al. (2006) A sodium transporter (HKT7) is a candidate for *Nax1*, a gene for salt tolerance in durum wheat. *Plant Physiol.* 142, 1718–1727.

- Imai, R., Koike, M., Sutoh, K., Kawakami, A., Torada, A. and Oono, K. (2005) Molecular characterization of a cold-induced plasma membrane protein gene from wheat. *Mol. Genet. Genomics* 274, 445–453.
- In, O., Berberich, T., Romdhane, S. and Feierabend, J. (2005) Changes in gene expression during dehardening of cold-hardened winter rye (*Secale cereale* L.) leaves and potential role of a peptide methionine sulfoxide reductase in cold-acclimation. *Planta* 220, 941–950.
- Ito, K., Tanakamaru, K., Morita, S., Abe, J. and Inanaga, S. (2006) Lateral root development, including responses to soil drying, of maize (*Zea mays*) and wheat (*Triticum aestivum*) seminal roots. *Physiol. Plantarum* 127, 260–267.
- Ivancic, V., Hackett, C.A., Zhang, Z.J., Staub, J.E., Nevo, E., Thomas, W.T.B., et al. (2000) Phenotypic responses of wild barley to experimentally imposed water stress. *J. Exp. Bot.* 51, 2021–2029.
- James, R.A., Davenport, R.J. and Munns, R. (2006) Physiological characterization of two genes for Na⁺ exclusion in durum wheat, *Nax1* and *Nax2*. *Plant Physiol.* 142, 1537–1547.
- James, V.A., Neibaur, I. and Altpeter, F. (2008) Stress inducible expression of the DREB1A transcription factor from xeric, *Hordeum spontaneum* L. in turf and forage grass (*Paspalum notatum* Flugge) enhances abiotic stress tolerance. *Transgenic Res.* 17, 93–104.
- Jefferies, S.P., Barr, A.R., Karakousis, A., Kretschmer, J.M., Manning, S., Chalmers, K.J., et al. (1999) Mapping of chromosome regions conferring boron toxicity tolerance in barley (*Hordeum vulgare* L.). *Theor. Appl. Genet.* 98, 1293–1303.
- Jefferies, S.P., Pallotta, M.A., Paull, J.G., Karakousis, A., Kretschmer, J.M., Manning, S., et al. (2000) Mapping and validation of chromosome regions conferring boron toxicity tolerance in wheat (*Triticum aestivum*). *Theor. Appl. Genet.* 101, 767–777.
- Jukanti, A.K., Heidlebaugh, N.M., Parrott, D.L., Fischer, I.A., McInerney, K. and Fischer, A.M. (2008) Comparative transcriptome profiling of near-isogenic barley (*Hordeum vulgare*) lines differing in the allelic state of a major grain protein content locus identifies genes with possible roles in leaf senescence and nitrogen reallocation. *New Phytol.* 177, 333–349.
- Jung, J., Won, S.Y., Suh, S.C., Kim, H., Wing, R., Jeong, Y., et al. (2007) The barley ERF-type transcription factor HvRAF confers enhanced pathogen resistance and salt tolerance in Arabidopsis. *Planta* 225, 575–588.
- Kalayci, M., Alkan, A., Cakmak, I., Bayramoglu, O., Yilmaz, A., Aydin, M., et al. (1998) Studies on differential response of wheat cultivars to boron toxicity (Reprinted from wheat: prospects for global improvement, 1998). *Euphytica* 100, 123–129.
- Karakousis, A., Barr, A.R., Kretschmer, J.M., Manning, S., Jefferies, S.P., Chalmers, K.J., et al. (2003) Mapping and QTL analysis of the barley population Clipper × Sahara. *Aust. J. Agric. Res.* 54, 1137–1140.
- Katsuhara, M. and Shibasaka, M. (2007) Barley root hydraulic conductivity and aquaporins expression in relation to salt tolerance. *Soil Sci. Plant Nutr.* 53, 466–470.
- Kawaura, K., Mochida, K., Yamazaki, Y. and Ogiwara, Y. (2006) Transcriptome analysis of salinity stress responses in common wheat using a 22 k oligo-DNA microarray. *Funct. Integr. Genomics* 6, 132–142.
- Khan, P., Hoffmann, L., Renaut, J. and Hausman, J.F. (2007) Current initiatives in proteomics for the analysis of plant salt tolerance. *Curr. Sci.* 93, 807–817.
- Kilian, A. (2005) The fast and the cheap: SNP and DArT-based whole genome profiling for crop improvement. In: R. Tuberosa, R.L. Phillips and M. Gale (Eds.), *Proceedings of an International Congress "In the Wake of the Double Helix: From the Green Revolution to the Gene Revolution"*, 27–31 May 2003, Bologna, Italy. Avenue media, Bologna, pp. 443–461.
- Kimurto, P.K., Kinyua, M.G., Birechl, R., Korir, P.C., Njoka, E.M. and Njau, P.N. (2005) Root and shoot characteristics as selection criteria for drought tolerance in bread wheat (*Triticum aestivum* L.) at seedling stage under tropical environment. *Discov. Innovat.* 17, 74–84.

- Kirigwi, F.M., Van Ginkel, M., Brown-Guedira, G., Gill, B.S., Paulsen, G.M. and Fritz, A.K. (2007) Markers associated with a QTL for grain yield in wheat under drought. *Mol. Breed.* 20, 401–413.
- Kirst, M. and Yu, Q. (2007) Genetical genomics: successes and prospect in plants. In: R.K. Varshney and R. Tuberosa (Eds.), *Genomics-Assisted Crop Improvement-Vol 1: Genomics Approaches and Platforms*. Springer, Dordrecht, pp. 245–265.
- Knox, A.K., Li, C.X., Vagujfalvi, A., Galilba, G., Stockinger E.J. and Dubcovsky, J. (2008) Identification of candidate CBF genes for the frost tolerance locus Fr-Am2 in *Triticum monococcum*. *Plant Mol. Biol.* 67, 257–270.
- Kobayashi, F., Takumi, S., Nakata, M., Ohno, R., Nakamura, T. and Nakamura, C. (2004) Comparative study of the expression profiles of the *Cor/Lea* gene family in two wheat cultivars with contrasting levels of freezing tolerance. *Physiol. Plantarum* 120, 585–594.
- Kobayashi, F., Takumi, S., Kume, S., Ishibashi, M., Ohno, R., Murai, K., et al. (2005) Regulation by *Vrn-1/Fr-1* chromosomal intervals of CBF-mediated *Cor/Lea* gene expression and freezing tolerance in common wheat. *J. Exp. Bot.* 56, 887–895.
- Kochian, L.V., Pineros, M.A. and Hoekenga, O.A. (2005) The physiology, genetics and molecular biology of plant aluminum resistance and toxicity. *Plant Soil* 274, 175–195.
- Kubo, K., Elouafi, L., Watanabe, N., Nachit, M.M., Inagaki, M.N., Iwama, K., et al. (2007) Quantitative trait loci for soil-penetrating ability of roots in durum wheat. *Plant Breed.* 126, 375–378.
- Kuchel, H., Ye, G.Y., Fox, R. and Jefferies, S. (2005) Genetic and economic analysis of a targeted marker-assisted wheat breeding strategy. *Mol. Breed.* 16, 67–78.
- Kuchel, H., Fox, R., Reinheimer, J., Mosionek, L., Willey, N., Bariana, H., et al. (2007a) The successful application of a marker-assisted wheat breeding strategy. *Mol. Breed.* 20, 295–308.
- Kuchel, H., Williams, K., Langridge, P., Eagles, H.A. and Jefferies, S.P. (2007b) Genetic dissection of grain yield in bread wheat. II. QTL-by-environment interaction. *Theor. Appl. Genet.* 115, 1015–1027.
- Kuchel, H., Williams, K.J., Langridge, P., Eagles, H.A. and Jefferies, S.P. (2007c) Genetic dissection of grain yield in bread wheat. I. QTL analysis. *Theor. Appl. Genet.* 115, 1029–1041.
- Kume, S., Kobayashi, F., Ishibashi, M., Ohno, R., Nakamura, C. and Takumi, S. (2005) Differential and coordinated expression of *Cbf* and *Cor/Lea* genes during long-term cold acclimation in two wheat cultivars showing distinct levels of freezing tolerance. *Genes Genet. Syst.* 80, 185–197.
- Landjeva, S., Neumann, K., Lohwasser, U. and Börner, A. (2008) Molecular mapping of genomic regions associated with wheat seedling growth under osmotic stress. *Biol. Plantarum* 52, 259–266.
- Langridge, P. (2005) Molecular breeding of wheat and barley. In: R. Tuberosa, R.L. Phillips and M. Gale (Eds.), *Proceedings of the International Congress "In the Wake of the Double Helix: From the Green Revolution to the Gene Revolution"*, 27–31 May 2003, Bologna, Italy. Avenue media, Bologna, pp. 279–286.
- Laperche, A., Devienne-Barret, F., Mauray, O., Le Gouis, J. and Ney, B. (2006) A simplified conceptual model of carbon/nitrogen functioning for QTL analysis of winter wheat adaptation to nitrogen deficiency. *Theor. Appl. Genet.* 113, 1131–1146.
- Laperche, A., Brancourt-Hulmel, M., Heumez, E., Gardet, O., Hanocq, E., Devienne-Barret, F., et al. (2007) Using genotype × nitrogen interaction variables to evaluate the QTL involved in wheat tolerance to nitrogen constraints. *Theor. Appl. Genet.* 115, 399–415.
- Laurie, D.A., Pratchett, N., Bezant, J.H. and Snape, J.W. (1995) RFLP mapping of five major genes and eight QTL controlling flowering time in a winter × spring barley (*Hordeum vulgare* L.) cross. *Genome* 38, 575–585.
- Lee, M. (1995) DNA markers and plant breeding programs. In: D.L. Sparks (Ed.), *Advances in Agronomy Vol 55*. Academic Press Inc., San Diego, pp. 265–344.

- Lee, T.G., Jang, C.S., Kim, J.Y., Seong, R.C., Kim, I.G., Kim, D.S., et al. (2007) Expressed sequence tags from wheat roots under hypoxia. *Russ. J. Plant Physiol.* 54, 659–668.
- Levitt, J. (1972) *Responses of Plants to Environmental Stresses*. Academic Press, New York.
- Li, A.L., Zhu, Y.F., Tan, X.M., Wang, X., Wei, B., Guo, H.Z., et al. (2008) Evolutionary and functional study of the CDPK gene family in wheat (*Triticum aestivum* L.). *Plant Mol. Biol.* 66, 429–443.
- Li, C.D., Zhang, G.P. and Lance, R. (2007) Recent advances in breeding barley for drought and saline stress tolerance. In: M.A. Jenks, P.M. Hasegawa and S. Mohan Jain (Eds.), *Advances in Molecular Breeding Toward Drought and Salt Tolerant Crops*. Springer, Dordrecht, pp. 603–626.
- Liang, C.Z., Jaiswal, P., Hebbard, C., Avraham, S., Buckler, E.S., Casstevens, T., et al. (2008) Gramene: a growing plant comparative genomics resource. *Nucl. Acids Res.* 36, D947–D953.
- Lilley, J.M., Ludlow, M.M., McCouch, S.R. and Otoole, J.C. (1996) Locating QTL for osmotic adjustment and dehydration tolerance in rice. *J. Exp. Bot.* 47, 1427–1436.
- Lim, K.J., Junttila, S., Fey, V. and Rudd, S. (2007) Genomeless genomics in crop improvement. In: R.K. Varshney and R. Tuberosa (Eds.), *Genomics-Assisted Crop Improvement-Vol 1: Genomics Approaches and Platforms*. Springer, Dordrecht, pp. 151–176.
- Limin, A.E. and Fowler, D.B. (2006) Low-temperature tolerance and genetic potential in wheat (*Triticum aestivum* L.): response to photoperiod, vernalization, and plant development. *Planta* 224, 360–366.
- Lindsay, M.P., Lagudah, E.S., Hare, R.A. and Munns, R. (2004) A locus for sodium exclusion (*Nax1*), a trait for salt tolerance, mapped in durum wheat. *Funct. Plant Biol.* 31, 1105–1114.
- Ludlow, M.M. and Muchow, R.C. (1990) A critical evaluation of traits for improving crop yields in water-limited environments. *Adv. Agron.* 43, 107–153.
- Lukens, L.N. and Zhan, S.H. (2007) The plant genome's methylation status and response to stress: implications for plant improvement. *Curr. Opin. Plant Biol.* 10, 317–322.
- Ma, J.F., Shen, R.F., Nagao, S. and Tanimoto, E. (2004) Aluminum targets elongating cells by reducing cell wall extensibility in wheat roots. *Plant Cell Physiol.* 45, 583–589.
- Ma, L.Q., Zhou, E.F., Huo, N.X., Zhou, R.H., Wang, G.Y. and Jia, J.Z. (2007) Genetic analysis of salt tolerance in a recombinant inbred population of wheat (*Triticum aestivum* L.). *Euphytica* 153, 109–117.
- Maccaferri, M., Sanguineti, M.C. and Tuberosa, R. (2005) Analysis of linkage disequilibrium in a collection of elite durum wheat genotypes. *Mol. Breed.* 15, 271–289.
- Maccaferri, M., Sanguineti, M.C., Corneti, S., Ortega, J.L.A., Ben Salem, M., Bort, J., et al. (2008) Quantitative trait loci for grain yield and adaptation of durum wheat (*Triticum durum* Desf.) across a wide range of water availability. *Genetics* 178, 489–511.
- Magalhaes, J.V., Liu, J., Guimaraes, C.T., Lana, U.G.P., Alves, V.M.C., Wang, Y.H., et al. (2007) A gene in the multidrug and toxic compound extrusion (MATE) family confers aluminum tolerance in sorghum. *Nat. Genet.* 39, 1156–1161.
- Mano, Y. and Takeda, K. (1997) Mapping quantitative trait loci for salt tolerance at germination and at the seedling stage of barley (*Hordeum vulgare* L.). *Euphytica* 94, 263–272.
- Manske, G.G.B., Ortiz-Monasterio, J.I., van Ginkel, R.M., Rajaram, S. and Vlek, P.L.G. (2002) Phosphorus use efficiency in tall, semi-dwarf and dwarf near-isogenic lines of spring wheat. *Euphytica* 125, 113–119.
- Mantovani, P., Maccaferri, M., Sanguineti, M.C., Tuberosa, R., Catizone, I., Wenzl, P., et al. (2008) An integrated DArT-SSR linkage map of durum wheat. *Mol. Breed.* 22, 629–648.
- Marè, C., Mazzucotelli, E., Crosatti, C., Francia, E., Stanca, A.M. and Cattivelli, L. (2004) Hv-WRKY38: a new transcription factor involved in cold- and drought-response in barley. *Plant Mol. Biol.* 55, 399–416.

- Martinez-Noel, G., Nagaraj, V.J., Calo, G., Wiemken, A. and Pontis, H.G. (2007) Sucrose regulated expression of a Ca^{2+} -dependent protein kinase (TaCDPK1) gene in excised leaves of wheat. *Plant Physiol. Biochem.* 45, 410–419.
- Matos, M., Camacho, M.V., Perez-Flores, V., Pernaute, B., Pinto-Carnide, O. and Benito, C. (2005) A new aluminum tolerance gene located on rye chromosome arm 7RS. *Theor. Appl. Genet.* 111, 360–369.
- Matos, M., Perez-Flores, V., Camacho, M.V., Pernaute, B., Pinto-Carnide, O. and Benito, C. (2007) Detection and mapping of SSRs in rye ESTs from aluminium-stressed roots. *Mol. Breed.* 20, 103–115.
- Mickelson, S., See, D., Meyer, F.D., Garner, J.P., Foster, C.R., Blake, T.K., et al. (2003) Mapping of QTL associated with nitrogen storage and remobilization in barley (*Hordeum vulgare* L.) leaves. *J. Exp. Bot.* 54, 801–812.
- Mifflin, B.J. and Habash, D.Z. (2002) The role of glutamine synthetase and glutamate dehydrogenase in nitrogen assimilation and possibilities for improvement in the nitrogen utilization of crops. *J. Exp. Bot.* 53, 979–987.
- Miftahudin, R., Scoles, G.J. and Gustafson, J.P. (2002) AFLP markers tightly linked to the aluminum-tolerance gene *Alt3* in rye (*Secale cereale* L.). *Theor. Appl. Genet.* 104, 626–631.
- Miftahudin, T., Chikmawati, T., Ross, K., Scoles, G.J. and Gustafson, J.P. (2005) Targeting the aluminum tolerance gene *Alt3* region in rye, using rice/rye micro-colinearity. *Theor. Appl. Genet.* 110, 906–913.
- Milla, M.A.R. and Gustafson, J.P. (2001) Genetic and physical characterization of chromosome 4DL in wheat. *Genome* 44, 883–892.
- Miller, A.K., Galiba, G. and Dubcovsky, J. (2006) A cluster of 11 *CBF* transcription factors is located at the frost tolerance locus *Fr-A''2* in *Triticum monococcum*. *Mol. Genet. Genomics* 275, 193–203.
- Mohammadi, M., Kav, N.N.V. and Deyholos, M.K. (2007) Transcriptional profiling of hexaploid wheat (*Triticum aestivum* L.) roots identifies novel, dehydration-responsive genes. *Plant Cell Environ.* 30, 630–645.
- Monroy, A.F., Dryanova, A., Malette, B., Oren, D.H., Farajalla, M.R., Liu, W., et al. (2007) Regulatory gene candidates and gene expression analysis of cold acclimation in winter and spring wheat. *Plant Mol. Biol.* 64, 409–423.
- Morgan, J.M. (1991) A gene controlling differences in osmoregulation in wheat. *Aust. J. Plant Physiol.* 18, 249–257.
- Morgenthal, K., Weckwerth, W. and Steuer, R. (2006) Metabolomic networks in plants: transitions from pattern recognition to biological interpretation. *Biosystems* 83, 108–117.
- Mott, I.W. and Wang, R.R.C. (2007) Comparative transcriptome analysis of salt-tolerant wheat germplasm lines using wheat genome arrays. *Plant Sci.* 173, 327–339.
- Mullan, D.J., Colmer, T.D. and Francki, M.G. (2007) Arabidopsis-rice-wheat gene orthologues for Na^+ transport and transcript analysis in wheat-*L. elongatum* aneuploids under salt stress. *Mol. Genet. Genomics* 277, 199–212.
- Munns, R. (1993) Physiological processes limiting plant-growth in saline soils: some dogmas and hypotheses. *Plant Cell Environ.* 16, 15–24.
- Munns, R., Rebetzke, G.J., Husain, S., James, R.A. and Hare, R.A. (2003) Genetic control of sodium exclusion in durum wheat. *Aust. J. Agric. Res.* 54, 627–635.
- Munns, R. (2005) Genes and salt tolerance: bringing them together. *New Phytol.* 167, 645–663.
- Munns, R., James, R.A. and Lauchli, A. (2006) Approaches to increasing the salt tolerance of wheat and other cereals. *J. Exp. Bot.* 57, 1025–1043.
- Munns, R. (2007) Prophylactically parking sodium in the plant. *New Phytol.* 176, 501–504.
- Munns, R. and Richards, R.A. (2007) Recent advances in breeding wheat for drought and salt stresses. In: M.A. Jenks, P.M. Hasegawa and S. Mohan Jain (Eds.), *Advances in Molecular Breeding Toward Drought and Salt Tolerant Crops*. Springer, Dordrecht, pp. 565–585.

- Nakazono, M., Qiu, F., Borsuk, L.A. and Schnable, P.S. (2003) Laser-capture microdissection, a tool for the global analysis of gene expression in specific plant cell types: identification of genes expressed differentially in epidermal cells or vascular tissues of maize. *Plant Cell* 15, 583–596.
- Negishi, T., Nakanishi, H., Yazaki, J., Kishimoto, N., Fujii, F., Shimbo, K., et al. (2002) cDNA microarray analysis of gene expression during Fe-deficiency stress in barley suggests that polar transport of vesicles is implicated in phytosiderophore secretion in Fe-deficient barley roots. *Plant J.* 30, 83–94.
- Nelson, T., Tausta, S.L., Gandotra, N. and Liu, T. (2006) Laser microdissection of plant tissue: what you see is what you get. *Annu. Rev. Plant Biol.* 57, 181–201.
- Nguyen, H.T. and Blum, A. (2004) *Physiology and Biotechnology Integration for Plant Breeding*. Marcel Dekker Inc., New York.
- Ogo, Y., Nakanishi Itai, R.N., Nakanishi, H., Inoue, H., Kobayashi, T., Suzuki, M., et al. (2006) Isolation and characterization of IRO2, a novel iron-regulated bHLH transcription factor in graminaceous plants. *J. Exp. Bot.* 57, 2867–2878.
- Oh, S.J., Kwon, C.W., Choi, D.W., Song, S.I. and Kim, J.K. (2007) Expression of barley *HvCBF4* enhances tolerance to abiotic stress in transgenic rice. *Plant Biotechnol. J.* 5, 646–656.
- Ohno, R., Takumi, S. and Nakamura, C. (2001) Expression of a cold-responsive *Lt-Cor* gene and development of freezing tolerance during cold acclimation in wheat (*Triticum aestivum* L.). *J. Exp. Bot.* 52, 2367–2374.
- Omielan, J.A., Epstein, E. and Dvorak, J. (1991) Salt tolerance and ionic relations of wheat as affected by individual chromosomes of salt-tolerant *Lophopyrum elongatum*. *Genome* 34, 961–974.
- Oraby, H.F., Ransom, C.B., Kravchenko, A.N. and Sticklen, M.B. (2005) Barley *HVA1* gene confers salt tolerance in R3 transgenic oat. *Crop Sci.* 45, 2218–2227.
- Oshino, T., Abiko, M., Saito, R., Ichiishi, E., Endo, M., Kawagishi-Kobayashi, M., et al. (2007) Premature progression of anther early developmental programs accompanied by comprehensive alterations in transcription during high-temperature injury in barley plants. *Mol. Genet. Genomics* 278, 31–42.
- Ozturk, Z.N., Talamè, V., Deyholos, M., Michalowski, C.B., Galbraith, D.W., Gozukirmizi, N., et al. (2002) Monitoring large-scale changes in transcript abundance in drought- and salt-stressed barley. *Plant Mol. Biol.* 48, 551–573.
- Pakniyat, H., Powell, W., Baird, E., Handley, L.L., Robinson, D., Scrimgeour, C.M., et al. (1997) AFLP variation in wild barley (*Hordeum spontaneum* C. Koch) with reference to salt tolerance and associated ecogeography. *Genome* 40, 332–341.
- Pallotta, M.A., Graham, R.D., Langridge, P., Sparrow, D.H.B. and Barker, S.J. (2000) RFLP mapping of manganese efficiency in barley. *Theor. Appl. Genet.* 101, 1100–1108.
- Pan, A., Hayes, P.M., Chen, F., Chen, T.H.H., Blake, T., Wright, S., et al. (1994) Genetic analysis of the components of winterhardiness in barley (*Hordeum vulgare* L.). *Theor. Appl. Genet.* 89, 900–910.
- Passioura, J. (2002) Environmental biology and crop improvement. *Funct. Plant Biol.* 29, 537–546.
- Passioura, J. (2007) The drought environment: physical, biological and agricultural perspectives. *J. Exp. Bot.* 58, 113–117.
- Patterson, J., Ford, K., Cassin, A., Natera, S. and Bacic, A. (2007) Increased abundance of proteins involved in phytosiderophore production in boron tolerant barley. *Plant Physiol.* 144, 1612–1631.
- Peighambari, S.A., Samadi, B.Y., Nabipour, A., Charmet, G. and Sarrafi, A. (2005) QTL analysis for agronomic traits in a barley doubled haploids population grown in Iran. *Plant Sci.* 169, 1008–1013.
- Pfluger, J. and Wagner, D. (2007) Histone modifications and dynamic regulation of genome accessibility in plants. *Curr. Opin. Plant Biol.* 10, 645–652.

- Philipps, G., Drzewiecki, C., Barth, O., Zschiesche, W. and Humbeck, K. (2006) Light-dependent expression of the cold-regulated gene *HvMCI* in barley (*Hordeum vulgare* L.). *J. Therm. Biol.* 31, 473–482.
- Plaschke, J., Borner, A., Xie, D.X., Koeber, R.M.D., Schlegel, R. and Gale, M.D. (1993) RFLP mapping of genes affecting plant height and growth habit in rye. *Theor. Appl. Genet.* 85, 1049–1054.
- Pop, M. and Salzberg, S.L. (2008) Bioinformatics challenges of new sequencing technology. *Trends Genet.* 24, 142–149.
- Potokina, E., Druka, A., Luo, Z.W., Wise, R., Waugh, R. and Kearsey, M. (2008) Gene expression quantitative trait locus analysis of 16,000 barley genes reveals a complex pattern of genome-wide transcriptional regulation. *Plant J.* 53, 90–101.
- Quarrie, S.A. (1991) Implications of genetic differences in ABA accumulation for crop production. In: W.J. Davies and H.G. Jones (Eds.), *Abscisic Acid: Physiology and Biochemistry*. BIOS Scientific Publishers, Oxford, pp. 227–243.
- Quarrie, S.A., Gulli, M., Calestani, C., Steed, A. and Marmioli, N. (1994a) Location of a gene regulating drought-induced abscisic-acid production on the long arm of chromosome 5a of wheat. *Theor. Appl. Genet.* 89, 794–800.
- Quarrie, S.A., Lebreton, C., Gulli, M., Calestani, C. and Marmioli, N. (1994b) QTL analysis of ABA production in wheat and maize and associated physiological traits. *Russ. J. Plant Physiol.* 41, 565–571.
- Quarrie, S.A., Steed, A., Calestani, C., Semikhodskii, A., Lebreton, C., Chinoy, C., et al. (2005) A high-density genetic map of hexaploid wheat (*Triticum aestivum* L.) from the cross Chinese Spring \times SQ1 and its use to compare QTLs for grain yield across a range of environments. *Theor. Appl. Genet.* 110, 865–880.
- Rafalski, A. and Morgante, M. (2004) Corn and humans: recombination and linkage disequilibrium in two genomes of similar size. *Trends Genet.* 20, 103–111.
- Raman, H., Moroni, J.S., Sato, K., Read, B.J. and Scott, B.J. (2002) Identification of AFLP and microsatellite markers linked with an aluminium tolerance gene in barley (*Hordeum vulgare* L.). *Theor. Appl. Genet.* 105, 458–464.
- Raman, H., Karakousis, A., Moroni, J.S., Raman, R., Read, B.J., Garvin, D.F., et al. (2003) Development and allele diversity of microsatellite markers linked to the aluminium tolerance gene *Alp* in barley. *Aust. J. Agric. Res.* 54, 1315–1321.
- Raman, H., Zhang, K.R., Cakir, M., Appels, R., Garvin, D.F., Maron, L.G., et al. (2005) Molecular characterization and mapping of *ALMT1*, the aluminium-tolerance gene of bread wheat (*Triticum aestivum* L.). *Genome* 48, 781–791.
- Raman, H., Raman, R., Wood, R. and Martin, P. (2006) Repetitive indel markers within the *ALMT1* gene conditioning aluminium tolerance in wheat (*Triticum aestivum* L.). *Mol. Breed.* 18, 171–183.
- Raman, H., Ryan, P.R., Raman, R., Stodart, B.J., Zhang, K., Martin, P., et al. (2008) Analysis of *TaALMT1* traces the transmission of aluminum resistance in cultivated common wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* 116, 343–354.
- Rao, S.A. and McNeilly, T. (1999) Genetic basis of variation for salt tolerance in maize (*Zea mays* L.). *Euphytica* 108, 145–150.
- Rebetzke, G.J., van Herwaarden, A.F., Jenkins, C., Weiss, M., Lewis, D., Ruuska, S., et al. (2008) Quantitative trait loci for water-soluble carbohydrates and associations with agronomic traits in wheat. *Austr. J. Agric. Res.* 10, 891–905.
- Reinheimer, J.L., Barr, A.R. and Eglinton, J.K. (2004) QTL mapping of chromosomal regions conferring reproductive frost tolerance in barley (*Hordeum vulgare* L.). *Theor. Appl. Genet.* 109, 1267–1274.
- Ren, Z.H., Gao, J.P., Li, L.G., Cai, X.L., Huang, W., Chao, D.Y., et al. (2005) A rice quantitative trait locus for salt tolerance encodes a sodium transporter. *Nat. Genet.* 37, 1141–1146.

- Rerkasem, B., Nirantrayagul, S. and Jamjod, S. (2004) Increasing boron efficiency in international bread wheat, durum wheat, triticale and barley germplasm will boost production on soils low in boron. *Field Crops Res.* 86, 175–184.
- Reynolds, M.P., Mujeeb-Kazi, A. and Sawkins, M. (2005) Prospects for utilising plant-adaptive mechanisms to improve wheat and other crops in drought- and salinity-prone environments. *Ann. Appl. Biol.* 146, 239–259.
- Reynolds, M.P., Dreccer, F. and Trethowan, R. (2007a) Drought-adaptive traits derived from wheat wild relatives and landraces. *J. Exp. Bot.* 58, 177–186.
- Reynolds, M.P., Braun, H.J., Pietragalla, J. and Ortiz, R. (2007b). Challenges to international wheat breeding. *Euphytica* 157, 281–285.
- Reynolds, M.P. and Tuberosa, R. (2008) Translational research impacting on crop productivity in drought-prone environments. *Curr. Opin. Plant Biol.* 11, 171–179.
- Ribaut, J.M. and Hoisington, D. (1998) Marker-assisted selection: new tools and strategies. *Trends Plant Sci.* 3, 236–239.
- Richards, R.A., Dennet, C.W., Qualset, C.O., Epstein, E., Norlyn, J.D. and Winslow, M.D. (1987) Variation in yield and biomass in wheat, barley, and triticale in a salt-affected field. *Field Crops Res.* 15, 277–287.
- Richards, R.A. (1996) Defining selection criteria to improve yield under drought. *Plant Growth Regul.* 20, 157–166.
- Richards, R.A. (2006) Physiological traits used in the breeding of new cultivars for water-scarce environments. *Agr. Water Manage.* 80, 197–211.
- Richards, R.A., Watt, M. and Rebetzke, G.J. (2007) Physiological traits and cereal germplasm for sustainable agricultural systems. *Euphytica* 154, 409–425.
- Robinson, D., Handley, L.L., Scrimgeour, C.M., Gordon, D.C., Forster, B.P. and Ellis, R.P. (2000) Using stable isotope natural abundances ($\delta^{15}\text{N}$ and $\delta^{13}\text{C}$) to integrate the stress responses of wild barley (*Hordeum spontaneum* C. Koch) genotypes. *J. Exp. Bot.* 51, 41–50.
- Rodriguez, E.M., Svensson, J.T., Malatrasi, M., Choi, D.W. and Close, T.J. (2005) Barley *Dhn13* encodes a KS-type dehydrin with constitutive and stress responsive expression. *Theor. Appl. Genet.* 110, 852–858.
- Roessner, U., Patterson, J.H., Forbes, M.G., Fincher, G.B., Langridge, P. and Bacic, A. (2006) An investigation of boron toxicity in barley using metabolomics. *Plant Physiol.* 142, 1087–1101.
- Rostoks, N., Mudie, S., Cardle, L., Russell, J., Ramsay, L., Booth, A., et al. (2005) Genome-wide SNP discovery and linkage analysis in barley based on genes responsive to abiotic stress. *Mol. Genet. Genomics* 274, 515–527.
- Rostoks, N., Ramsay, L., MacKenzie, K., Cardle, L., Bhat, P.R., Roose, M.L., et al. (2006) Recent history of artificial outcrossing facilitates whole-genome association mapping in elite inbred crop varieties. *Proc. Natl. Acad. Sci. USA* 103, 18656–18661.
- Rus, A., Baxter, I., Muthukumar, B., Gustin, J., Lahner, B., Yakubova, E., et al. (2006) Natural variants of *AtHKT1* enhance Na^+ accumulation in two wild populations of *Arabidopsis*. *PLoS Genet.* 2, 1964–1973.
- Ruuska, S.A., Lewis, D.C., Kennedy, G., Furbank, R.T., Jenkins, C.L.D. and Tabe, L.M. (2008) Large scale transcriptome analysis of the effects of nitrogen nutrition on accumulation of stem carbohydrate reserves in reproductive stage wheat. *Plant Mol. Biol.* 66, 15–32.
- Sadok, W., Naudin, P., Boussuge, B., Muller, B., Welcker, C. and Tardieu, F. (2007) Leaf growth rate per unit thermal time follows QTL-dependent daily patterns in hundreds of maize lines under naturally fluctuating conditions. *Plant Cell Environ.* 30, 135–146.
- Saini, H.S. and Westgate, M.E. (2000) Reproductive development in grain crops during drought. *Adv. Agron.* 68, 59–96.
- Sakurai, N. and Shibata, D. (2006) KaPPA-view for integrating quantitative transcriptomic and metabolomic data on plant metabolic pathway maps. *J. Pestic. Sci.* 31, 293–295.

- Salem, K.F.M., Roder, M.S. and Borner, A. (2007) Identification and mapping quantitative trait loci for stem reserve mobilisation in wheat (*Triticum aestivum* L.). *Cereal Res. Commun.* 35, 1367–1374.
- Salse, J. and Feuillet, C. (2007) Comparative genomics of cereals. In: R.K. Varshney and R. Tuberosa (Eds.), *Genomics-Assisted Crop Improvement-Vol 1: Genomics Approaches and Platforms*. Springer, Dordrecht, pp. 177–206.
- Salvi, S. and Tuberosa, R. (2005) To clone or not to clone plant QTLs: present and future challenges. *Trends Plant Sci.* 10, 297–304.
- Salvi, S. and Tuberosa, R. (2007) Cloning QTLs in plants. In: R.K. Varshney and R. Tuberosa (Eds.), *Genomics-Assisted Crop Improvement-Vol 1: Genomics Approaches and Platforms*. Springer, Dordrecht, pp. 207–226.
- Salvi, S., Sponza, G., Morgante, M., Tomes, D., Niu, X., Fengler, K.A., et al. (2007) Conserved noncoding genomic sequences associated with a flowering-time quantitative trait locus in maize. *Proc. Natl. Acad. Sci. USA* 104, 11376–11381.
- Sanguineti, M.C., Tuberosa, R., Stefanelli, S., Noli, E., Blake, T.K., Hayes, P.M., et al. (1994) Utilization of a recombinant inbred population to localize QTLs for abscisic acid content in leaves of drought stressed barley (*Hordeum vulgare* L.). *Russ. J. Plant Physiol.* 41, 572–576.
- Sanguineti, M.C., Li, S., Maccaferri, M., Corneti, S., Rotondo, F., Chiari, T., et al. (2007) Genetic dissection of seminal root architecture in elite durum wheat germplasm. *Ann. Appl. Biol.* 151, 291–305.
- Sarma, R.N., Gill, B.S., Sasaki, T., Galiba, G., Sutka, J., Laurie, D.A., et al. (1998) Comparative mapping of the wheat chromosome 5A *Vrn-A1* region with rice and its relationship to QTL for flowering time. *Theor. Appl. Genet.* 97, 103–109.
- Sasaki, T., Yamamoto, Y., Ezaki, B., Katsuhara, M., Ahn, S.J., Ryan, P.R., et al. (2004) A wheat gene encoding an aluminum-activated malate transporter. *Plant J.* 37, 645–653.
- Sasaki, T., Ryan, P.R., Delhaize, E., Hebb, D.M., Ogihara, Y., Kawaura, K., et al. (2006) Sequence upstream of the wheat (*Triticum aestivum* L.) *ALMT1* gene and its relationship to aluminum resistance. *Plant Cell Physiol.* 47, 1343–1354.
- Schachtman, D. and Liu, W. (1999) Molecular pieces to the puzzle of the interaction between potassium and sodium uptake in plants. *Trends Plant Sci.* 4, 281–287.
- Schnurbusch, T., Collins, N.C., Eastwood, R.F., Sutton, T., Jefferies, S.P., Langridge, P., et al. (2007) Fine mapping and targeted SNP survey using rice-wheat gene colinearity in the region of the Bo1 boron toxicity tolerance locus of bread wheat. *Theor. Appl. Genet.* 115, 451–461.
- Schunmann, P.H.D., Richardson, A.E., Smith, F.W. and Delhaize, E. (2004a) Characterization of promoter expression patterns derived from the Pht1 phosphate transporter genes of barley (*Hordeum vulgare* L.). *J. Exp. Bot.* 55, 855–865.
- Schunmann, P.H.D., Richardson, A.E., Vickers, C.E. and Delhaize, E. (2004b) Promoter analysis of the barley Pht1;1 phosphate transporter gene identifies regions controlling root expression and responsiveness to phosphate deprivation. *Plant Physiol.* 136, 4205–4214.
- Senthil-Kumar, M., Kumar, G., Srikanthbabu, V. and Udayakumar, M. (2007) Assessment of variability in acquired thermotolerance: Potential option to study genotypic response and the relevance of stress genes. *J. Plant Physiol.* 164, 111–125.
- Setter, T.L. and Waters, I. (2003) Review of prospects for germplasm improvement for waterlogging tolerance in wheat, barley and oats. *Plant Soil* 253, 1–34.
- Sharp, R.E., Poroyko, V., Hejlek, L.G., Spollen, W.G., Springer, G.K., Bohnert, H.J., et al. (2004) Root growth maintenance during water deficits: physiology to functional genomics. *J. Exp. Bot.* 55, 2343–2351.
- Shen, Q.X., Uknes, S.J. and Ho, T.H.D. (1993) Hormone response complex in a novel abscisic acid and cycloheximide-inducible barley gene. *J. Biol. Chem.* 268, 23652–23660.
- Shen, Q.X.J., Casaretto, J.A., Zhang, P. and Ho, T.H.D. (2004) Functional definition of ABA-response complexes: the promoter units necessary and sufficient for ABA induction of gene expression in barley (*Hordeum vulgare* L.). *Plant Mol. Biol.* 54, 111–124.

- Shen, Y.G., Zhang, W.K., He, S.J., Zhang, J.S., Liu, Q. and Chen, S.Y. (2003) An EREBP/AP2-type protein in *Triticum aestivum* was a DRE-binding transcription factor induced by cold, dehydration and ABA stress. *Theor. Appl. Genet.* 106, 923–930.
- Shimamura, C., Ohno, R., Nakamura, C. and Takumi, S. (2006) Improvement of freezing tolerance in tobacco plants expressing a cold-responsive and chloroplast-targeting protein WCOR15 of wheat. *J. Plant Physiol.* 163, 213–219.
- Shinozaki, K. and Yamaguchi-Shinozaki, K. (2000) Molecular responses to dehydration and low temperature: differences and cross-talk between two stress signaling pathways. *Curr. Opin. Plant Biol.* 3, 217–223.
- Shinozaki, K. and Yamaguchi-Shinozaki, K. (2007) Gene networks involved in drought stress response and tolerance. *J. Exp. Bot.* 58, 221–227.
- Skinner, J.S., von Zitzewitz, J., Szucs, P., Marquez-Cedillo, L., Filichkin, T., Amundsen, K., et al. (2005) Structural, functional, and phylogenetic characterization of a large CBF gene family in barley. *Plant Mol. Biol.* 59, 533–551.
- Skinner, J.S., Szucs, P., von Zitzewitz, J., Marquez-Cedillo, L., Filichkin, T., Stockinger, E.J., et al. (2006) Mapping of barley homologs to genes that regulate low temperature tolerance in *Arabidopsis*. *Theor. Appl. Genet.* 112, 832–842.
- Slade, A.J. and Knauf, V.C. (2005) TILLING moves beyond functional genomics into crop improvement. *Transgenic Res.* 14, 109–115.
- Slade, A.J., Fuerstenberg, S.I., Loeffler, D., Steine, M.N. and Facciotti, D. (2005) A reverse genetic, nontransgenic approach to wheat crop improvement by TILLING. *Nat. Biotechnol.* 23, 75–81.
- Slavich, P.G., Read, B.J. and Cullis, B.R. (1990) Yield response of barley germplasm to field variation in salinity quantified using the EM-38. *Aust. J. Exp. Agric.* 30, 551–556.
- Snape, J.W., Semikhodskii, A., Fish, L., Sarma, R.N., Quarrie, S.A., Galiba, G., et al. (1997) Mapping frost tolerance loci in wheat and comparative mapping with other cereals. *Acta Agron. Hung.* 45, 268–270.
- Snape, J.W., Sarma, R., Quarrie, S.A., Fish, L., Galiba, G. and Sutka, J. (2001) Mapping genes for flowering time and frost tolerance in cereals using precise genetic stocks. *Euphytica* 120, 309–315.
- Snape, J.W., Foulkes, M.J., Simmonds, J., Leverington, M., Fish, L.J., Wang, Y., et al. (2007) Dissecting gene x environmental effects on wheat yields via QTL and physiological analysis. *Euphytica* 154, 401–408.
- Somers, D.J., Banks, T., DePauw, R., Fox, S., Clarke, J., Pozniak, C., et al. (2007) Genome-wide linkage disequilibrium analysis in bread wheat and durum wheat. *Genome* 50, 557–567.
- Song, X., Ni, Z.F., Yao, Y.Y., Xie, C.J., Li, Z.X., Wu, H.Y., et al. (2007) Wheat (*Triticum aestivum* L.) root proteome and differentially expressed root proteins between hybrid and parents. *Proteomics* 7, 3538–3557.
- Sørensen, A.P., Stuurman, J., Rouppe van der Voort, J. and Peleman, J. (2007) Molecular breeding: maximizing the exploitation of genetic diversity. In: R.K. Varshney and R. Tuberosa (Eds.), *Genomics-Assisted Crop Improvement-Vol 1: Genomics Approaches and Platforms*. Springer, Dordrecht, pp. 31–56.
- Sorrells, M.E., La Rota, M., Bermudez-Kandianis, C.E., Greene, R.A., Kantety, R., Munkvold, J.D., et al. (2003) Comparative DNA sequence analysis of wheat and rice genomes. *Genome Res.* 13, 1818–1827.
- Stein, N. (2007) Triticeae genomics: advances in sequence analysis of large genome cereal crops. *Chromosome Res.* 15, 21–31.
- Stein, N., Prasad, M., Scholz, U., Thiel, T., Zhang, H.N., Wolf, M., et al. (2007) A 1,000-loci transcript map of the barley genome: new anchoring points for integrative grass genomics. *Theor. Appl. Genet.* 114, 823–839.
- Stephenson, T.J., McIntyre, C.L., Collet, C. and Xue, G.P. (2007) Genome-wide identification and expression analysis of the NF-Y family of transcription factors in *Triticum aestivum*. *Plant Mol. Biol.* 65, 77–92.

- Stockinger, E.J., Skinner, J.S., Gardner, K.G., Francia, E. and Pecchioni, N. (2007) Expression levels of barley *Cbf* genes at the *Frost resistance-H2* locus are dependent upon alleles at *Fr-H1* and *Fr-H2*. *Plant J.* 51, 308–321.
- Su, J.Y., Tong, Y.P., Liu, Q.Y., Li, B., Jing, R.L., Li, J.Y., et al. (2006) Mapping quantitative trait loci for post-anthesis dry matter accumulation in wheat. *J. Integr. Plant Biol.* 48, 938–944.
- Sugie, A., Naydenov, N., Mizuno, N., Nakamura, C. and Takumi, S. (2006) Overexpression of wheat alternative oxidase gene *Waox1a* alters respiration capacity and response to reactive oxygen species under low temperature in transgenic Arabidopsis. *Genes Genet. Syst.* 81, 349–354.
- Sunarpi, Horie, T., Motoda, J., Kubo, M., Yang, H., Yoda, K., et al. (2005) Enhanced salt tolerance mediated by AtHKT1 transporter-induced Na⁺ unloading from xylem vessels to xylem parenchyma cells. *Plant J.* 44, 928–938.
- Sunkar, R., Chinnusamy, V., Zhu, J.H. and Zhu, J.K. (2007) Small RNAs as big players in plant abiotic stress responses and nutrient deprivation. *Trends Plant Sci.* 12, 301–309.
- Sutka, J. (1994) Genetic control of frost tolerance in wheat (*Triticum aestivum* L.). *Euphytica* 77, 277–282.
- Sutka, J., Galiba, G., Vagujfalvi, A., Gill, B.S. and Snape, J.W. (1999) Physical mapping of the *Vrn-A1* and *Fr1* genes on chromosome 5A of wheat using deletion lines. *Theor. Appl. Genet.* 99, 199–202.
- Sutton, T., Baumann, U., Hayes, J., Collins, N.C., Shi, B.J., Schnurbusch, T., et al. (2007) Boron-toxicity tolerance in barley arising from efflux transporter amplification. *Science* 318, 1446–1449.
- Suzuki, M., Takahashi, M., Tsukamoto, T., Watanabe, S., Matsuhashi, S., Yazaki, J., et al. (2006) Biosynthesis and secretion of mugineic acid family phytosiderophores in zinc-deficient barley. *Plant J.* 48, 85–97.
- Takumi, S., Koike, A., Nakata, M., Kume, S., Ohno, R. and Nakamura, C. (2003) Cold-specific and light-stimulated expression of a wheat (*Triticum aestivum* L.) *Cor* gene *Wcor15* encoding a chloroplast-targeted protein. *J. Exp. Bot.* 54, 2265–2274.
- Takumi, S., Shimamura, C. and Kobayashi, F. (2008) Increased freezing tolerance through up-regulation of downstream genes via the wheat CBF gene in transgenic tobacco. *Plant Physiol. Biochem.* 46, 205–211.
- Talamè, V., Sanguineti, M.C., Chiapparino, E., Bahri, H., Ben Salem, M., Forster, B.P., et al. (2004) Identification of *Hordeum spontaneum* QTL alleles improving field performance of barley grown under rainfed conditions. *Ann. Appl. Biol.* 144, 309–319.
- Talamè, V., Ozturk, N.Z., Bohnert, H.J. and Tuberosa, R. (2007) Barley transcript profiles under dehydration shock and drought stress treatments: a comparative analysis. *J. Exp. Bot.* 58, 229–240.
- Talamè, V., Bovina, R., Sanguineti, M.C., Tuberosa, R., Lundqvist, U. and Salvi, S. (2008) TILLMore, a resource for the discovery of chemically induced mutants in barley. *Plant Biotechnol. J.* 6, 477–485.
- Tambussi, E.A., Bort, J. and Araus, J.L. (2007) Water use efficiency in C-3 cereals under Mediterranean conditions: a review of physiological aspects. *Ann. Appl. Biol.* 150, 307–321.
- Tamura, K. and Yamada, T. (2007) A perennial ryegrass *CBF* gene cluster is located in a region predicted by conserved synteny between Poaceae species. *Theor. Appl. Genet.* 114, 273–283.
- Tanksley, S.D. (1993) Mapping polygenes. *Annu. Rev. Genet.* 27, 205–233.
- Tanksley, S.D. and Nelson, J.C. (1996) Advanced backcross QTL analysis: a method for the simultaneous discovery and transfer of valuable QTLs from unadapted germplasm into elite breeding lines. *Theor. Appl. Genet.* 92, 191–203.
- Tardieu, F. (2003) Virtual plants: modelling as a tool for the genomics of tolerance to water deficit. *Trends Plant Sci.* 8, 9–14.

- Teulat, B., Monneveux, P., Wery, J., Borries, C., Souyris, I., Charrier, A., et al. (1997) Relationships between relative water content and growth parameters under water stress in barley: a QTL study. *New Phytol.* 137, 99–107.
- Teulat, B., This, D., Khairallah, M., Borries, C., Ragot, C., Sourdille, P., et al. (1998) Several QTLs involved in osmotic adjustment trait variation in barley (*Hordeum vulgare* L.). *Theor. Appl. Genet.* 96, 688–698.
- Teulat, B., Borries, C. and This, D. (2001a) New QTLs identified for plant water status, water-soluble carbohydrate and osmotic adjustment in a barley population grown in a growth-chamber under two water regimes. *Theor. Appl. Genet.* 103, 161–170.
- Teulat, B., Merah, O., Souyris, I. and This, D. (2001b) QTLs for agronomic traits from a Mediterranean barley progeny grown in several environments. *Theor. Appl. Genet.* 103, 774–787.
- Teulat, B., Merah, O., Sirault, X., Borries, C., Waugh, R. and This, D. (2002) QTLs for grain carbon isotope discrimination in field-grown barley. *Theor. Appl. Genet.* 106, 118–126.
- Teulat, B., Zoumarou-Wallis, N., Rotter, B., Ben Salem, M., Bahri, H. and This, D. (2003) QTL for relative water content in field-grown barley and their stability across Mediterranean environments. *Theor. Appl. Genet.* 108, 181–188.
- Thomson, C.J., Armstrong, W., Waters, I. and Greenway, H. (1990) Aerenchyma formation and associated oxygen movement in seminal and nodal roots of wheat. *Plant Cell Environ.* 13, 395–403.
- Till, B.J., Comai, L. and Henikoff, S. (2007) TILLING and EcoTILLING for crop improvement. In: R.K. Varshney and R. Tuberosa (Eds.), *Genomics-Assisted Crop Improvement-Vol 1: Genomics Approaches and Platforms*. Springer, Dordrecht, pp. 333–350.
- Tittarelli, A., Milla, L., Vargas, F., Morales, A., Neupert, C., Meisel, L.A., et al. (2007) Isolation and comparative analysis of the wheat TaPT2 promoter: identification in silico of new putative regulatory motifs conserved between monocots and dicots. *J. Exp. Bot.* 58, 2573–2582.
- Tondelli, A., Francia, E., Barabaschi, D., Aprile, A., Skinner, J.S., Stockinger, E.J., et al. (2006) Mapping regulatory genes as candidates for cold and drought stress tolerance in barley. *Theor. Appl. Genet.* 112, 445–454.
- Tong, Y., Zhou, J.J., Li, Z.S. and Miller, A.J. (2005) A two-component high-affinity nitrate uptake system in barley. *Plant J.* 41, 442–450.
- Torres, T.T., Metta, M., Ottenwalder, B. and Schlotterer, C. (2008) Gene expression profiling by massively parallel sequencing. *Genome Res.* 18, 172–177.
- Toth, B., Galiba, G., Feher, E., Sutka, J. and Snape, J.W. (2003) Mapping genes affecting flowering time and frost resistance on chromosome 513 of wheat. *Theor. Appl. Genet.* 107, 509–514.
- Travella, S., Klimm, T.E. and Keller, B. (2006) RNA interference-based gene silencing as an efficient tool for functional genomics in hexaploid bread wheat. *Plant Physiol.* 142, 6–20.
- Trueman, L.J., Richardson, A. and Forde, B.G. (1996) Molecular cloning of higher plant homologues of the high-affinity nitrate transporters of *Chlamydomonas reinhardtii* and *Aspergillus nidulans*. *Gene* 175, 223–231.
- Tsuda, K., Tsvetanov, S., Takumi, S., Mori, N., Atanassov, A. and Nakamura, C. (2000) New members of a cold-responsive group-3 *Lea/Rab*-related *Cor* gene family from common wheat (*Triticum aestivum* L.). *Genes Genet. Syst.* 75, 179–188.
- Tsvetanov, S., Ohno, R., Tsuda, K., Takumi, S., Mori, N., Atanassov, A., et al. (2000) A cold-responsive wheat (*Triticum aestivum* L.) gene *wcor14* identified in a winter-hardy cultivar 'Mironovska 808'. *Genes Genet. Syst.* 75, 49–57.
- Tuberosa, R., Galiba, G., Sanguineti, M.C., Noli, E. and Sukta, J. (1997) Identification of QTL influencing freezing tolerance in barley. *Acta Agron. Hung.* 45, 413–417.
- Tuberosa, R., Gill, B.S. and Quarrie, S.A. (2002a) Cereal genomics: ushering in a brave new world. *Plant Mol. Biol.* 48, 445–449.

- Tuberosa, R., Salvi, S., Sanguineti, M.C., Landi, P., Maccaferri, M. and Conti, S. (2002b) Mapping QTLs regulating morpho-physiological traits and yield: case studies, shortcomings and perspectives in drought-stressed maize. *Ann. Bot.* 89, 941–963.
- Tuberosa, R., Sanguineti, M.C., Landi, P., Giuliani, M.M., Salvi, S. and Conti, S. (2002c) Identification of QTLs for root characteristics in maize grown in hydroponics and analysis of their overlap with QTLs for grain yield in the field at two water regimes. *Plant Mol. Biol.* 48, 697–712.
- Tuberosa, R. (2004) Molecular approaches to unravel the genetic basis of water use efficiency. In: M. Bacon (Ed.), *Water Use Efficiency in Plant Biology*. Blackwell, Oxford, pp. 228–301.
- Tuberosa, R. and Salvi, S. (2006) Genomics-based approaches to improve drought tolerance of crops. *Trends Plant Sci.* 11, 405–412.
- Tuberosa, R., Giuliani, S., Parry, M.A.J. and Araus, J.L. (2007) Improving water use efficiency in Mediterranean agriculture: what limits the adoption of new technologies? *Ann. Appl. Biol.* 150, 157–162.
- Turner, N.C. (1997) Further progress in crop water relations. *Adv. Agron.* 528, 293–338.
- Turpeinen, T., Tenhola, T., Manninen, O., Nevo, E. and Nissila, E. (2001) Microsatellite diversity associated with ecological factors in *Hordeum spontaneum* populations in Israel. *Mol. Ecol.* 10, 1577–1591.
- Tuteja, N. (2007) Mechanisms of high salinity tolerance in plants. *Method. Enzymol.* 428, 419–438.
- Uauy, C., Distelfeld, A., Fahima, T., Blechl, A. and Dubcovsky, J. (2006) Positional cloning of the wheat vernalization gene *VRN1*. *Science* 314, 1298–1301.
- Ueda, A., Kathiresan, A., Bennett, J. and Takabe, T. (2004) Comparative expression analyses using barley cDNA microarray in barley and rice under salt stress. *Plant Cell Physiol.* 45, S227.
- Ueda, A., Kathiresan, A., Bennett, J. and Takabe, T. (2006) Comparative transcriptome analyses of barley and rice under salt stress. *Theor. Appl. Genet.* 112, 1286–1294.
- Ueda, A., Yamamoto-Yamane, Y. and Takabe, T. (2007) Salt stress enhances proline utilization in the apical region of barley roots. *Biochem. Bioph. Res. Co.* 355, 61–66.
- Ugarte, C., Calderini, D.F. and Slafer, G.A. (2007) Grain weight and grain number responsiveness to pre-anthesis temperature in wheat, barley and triticale. *Field Crops Res.* 100, 240–248.
- Vagujfalvi, A., Galiba, G., Cattivelli, L. and Dubcovsky, J. (2003) The cold-regulated transcriptional activator *Chf3* is linked to the frost-tolerance locus *Fr-A2* on wheat chromosome 5A. *Mol. Genet. Genomics* 269, 60–67.
- Vagujfalvi, A., Aprile, A., Miller, A., Dubcovsky, J., Delugu, G., Galiba, G., et al. (2005) The expression of several *Chf* genes at the *Fr-A2* locus is linked to frost resistance in wheat. *Mol. Genet. Genomics* 274, 506–514.
- van Eeuwijk, F.A., Malosetti, M., Yin, X.Y., Struik, P.C. and Stam, P. (2005) Statistical models for genotype by environment data: from conventional ANOVA models to eco-physiological QTL models. *Austr. J. Agric. Res.* 56, 883–894.
- van Ginkel, M. and Ogonnaya, F. (2007) Novel genetic diversity from synthetic wheats in breeding cultivars for changing production conditions. *Field Crops Res.* 104, 86–94.
- Vargas, W.A., Pontis, H.G. and Salerno, G.L. (2007) Differential expression of alkaline and neutral invertases in response to environmental stresses: characterization of an alkaline isoform as a stress-response enzyme in wheat leaves. *Planta* 226, 1535–1545.
- Varshney, R.K., Graner, A. and Sorrells, M.E. (2005) Genomics-assisted breeding for crop improvement. *Trends Plant Sci.* 10, 621–630.
- Varshney, R.K. and Tuberosa, R. (2007) Genomics-assisted crop improvement: an overview. In: R.K. Varshney and R. Tuberosa (Eds.), *Genomics-Assisted Crop Improvement, Vol 1: Genomics Approaches and Platforms*. Springer, Dordrecht, pp. 1–12.

- Vasekina, A.V., Yershov, P.V., Reshetova, O.S., Tikhonova, T.V., Lunin, V.G., Trofimova, M.S., et al. (2005) Vacuolar Na^+/H^+ antiporter from barley: identification and response to salt stress. *Biochemistry (Moscow)* 70, 100–107.
- Verma, V., Foulkes, M.J., Worland, A.J., Sylvester-Bradley, R., Caligari, P.D.S. and Snape, J.W. (2004) Mapping quantitative trait loci for flag leaf senescence as a yield determinant in winter wheat under optimal and drought-stressed environments. *Euphytica* 135, 255–263.
- Veyrieras, J.B., Camus-Kulandaivelu, L., Gouesnard, B., Manicacci, D. and Charcosset, A. (2007) Bridging genomics and genetic diversity: linkage disequilibrium structure and association mapping in maize and other cereals. *Crop Sci.* 47, S60–S71.
- Vidmar, J.J., Zhuo, D.G., Siddiqi, M.Y. and Glass, A.D.M. (2000) Isolation and characterization of *HvNRT2.3* and *HvNRT2.4*, cDNAs encoding high-affinity nitrate transporters from roots of barley. *Plant Physiol.* 122, 783–792.
- Vij, S. and Tyagi, A.K. (2007) Emerging trends in the functional genomics of the abiotic stress response in crop plants. *Plant Biotechnol. J.* 5, 361–380.
- Walia, H., Wilson, C., Wahid, A., Condamine, P., Cui, X.P. and Close, T.J. (2005) Expression analysis of barley (*Hordeum vulgare* L.) during salinity stress. *Funct. Integr. Genomics* 6, 153–156.
- Walia, H., Wilson, C., Condamine, P., Ismail, A.M., Xu, J., Cui, X.P. and Close, T.J. (2007a) Array-based genotyping and expression analysis of barley cv. Maythorpe and Golden Promise. *BMC Genomics* 8, 87.
- Walia, H., Wilson, C., Condamine, P., Liu, X., Ismail, A.M. and Close T.J. (2007b) Large-scale expression profiling and physiological characterization of jasmonic acid-mediated adaptation of barley to salinity stress. *Plant Cell Environ.* 30, 410–421.
- Wang, J.P., Raman, H., Zhou, M.X., Ryan, P.R., Delhaize, E., Hebb, D.M., et al. (2007) High-resolution mapping of the *Alp* locus and identification of a candidate gene *HvMATE* controlling aluminium tolerance in barley (*Hordeum vulgare* L.). *Theor. Appl. Genet.* 115, 265–276.
- Wang, Y., Yao, J., Zhang, Z.F. and Zheng, Y.L. (2006) The comparative analysis based on maize integrated QTL map and meta-analysis of plant height QTLs. *Chinese Sci. Bull.* 51, 2219–2230.
- Waterhouse, P.M. and Helliwell, C.A. (2003) Exploring plant genomes by RNA-induced gene silencing. *Nat. Rev. Genet.* 4, 29–38.
- Way, H., Chapman, S., McIntyre, L., Casu, R., Xue, G.P., Manners, J., et al. (2005) Identification of differentially expressed genes in wheat undergoing gradual water deficit stress using a subtractive hybridisation approach. *Plant Sci.* 168, 661–670.
- Welcker, C., Boussuge, B., Bencivenni, C., Ribaut, J.M. and Tardieu, F. (2007) Are source and sink strengths genetically linked in maize plants subjected to water deficit? A QTL study of the responses of leaf growth and of Anthesis-Silking Interval to water deficit. *J. Exp. Bot.* 58, 339–349.
- Wight, C.P., Kibite, S., Tinker, N.A. and Molnar, S.J. (2006) Identification of molecular markers for aluminium tolerance in diploid oat through comparative mapping and QTL analysis. *Theor. Appl. Genet.* 112, 222–231.
- Wilkinson, M., McKibbin, R.S., Bailey, P.C., Flintham, J.E., Gale, M.D., Lenton, J.R., et al. (2002) Use of comparative molecular genetics to study pre harvest sprouting in wheat. *Euphytica* 126, 27–33.
- Wilkinson, M., Lenton, J. and Holdsworth, M. (2005) Transcripts of *Vp-1* homoeologues are alternatively spliced within the Triticeae tribe. *Euphytica* 143, 243–246.
- William, H.M., Trethowan, R. and Crosby-Galvan, E.M. (2007) Wheat breeding assisted by markers: CIMMYT's experience. *Euphytica* 157, 307–319.
- Wise, R.P., Moscou, M.J., Bogdanove, A.J. and Whitham, S.A. (2007) Transcript profiling in host-pathogen interactions. *Annu. Rev. Phytopathol.* 45, 329–369.
- Wissuwa, M., Ismail, A.M. and Yanagihara, S. (2006) Effects of zinc deficiency on rice growth and genetic factors contributing to tolerance. *Plant Physiol.* 142, 731–741.

- Witcombe, J.R., Hollington, P.A., Howarth, C.J., Reader, S. and Steele, K.A. (2008) Breeding for abiotic stresses for sustainable agriculture. *Philos. T. R. Soc. B* 363, 703–716.
- Wollenweber, B., Porter, J.R. and Lubberstedt, T. (2005) Need for multidisciplinary research towards a second green revolution. *Curr. Opin. Plant Biol.* 8, 337–341.
- Xia, L.Q., Ganai, M.W., Shewry, P.R., He, Z.H., Yang, Y. and Roder, M.S. (2008) Exploiting the diversity of *Viviparous-1* gene associated with pre-harvest sprouting tolerance in European wheat varieties. *Euphytica* 159, 411–417.
- Xiao, K., Bai, G.H. and Carver, B.F. (2005) Nylon filter arrays reveal differential expression of expressed sequence tags in wheat roots under aluminum stress. *J. Integr. Plant Biol.* 47, 839–848.
- Xiong, L. (2007) Absciscic acid in plant response and adaptation to drought and salt stress. In: M.A. Jenks, P.M. Hasegawa and S. Mohan Jain (Eds.), *Advances in Molecular Breeding Toward Drought and Salt Tolerant Crops*. Springer, Dordrecht, pp. 193–221.
- Xue, G.P. (2002) Characterisation of the DNA-binding profile of barley HvCBF1 using an enzymatic method for rapid, quantitative and high-throughput analysis of the DNA-binding activity. *Nucl. Acids Res.* 30, e-77.
- Xue, G.P. (2003) The DNA-binding activity of an AP2 transcriptional activator HvCBF2 involved in regulation of low-temperature responsive genes in barley is modulated by temperature. *Plant J.* 33, 373–383.
- Xue, G.P., McIntyre, C.L., Chapman, S., Bower, N.I., Way, H., Reverter, A., et al. (2006) Differential gene expression of wheat progeny with contrasting levels of transpiration efficiency. *Plant Mol. Biol.* 61, 863–881.
- Xue, Y., Jiang, L., Su, N., Wang, J.K., Deng, P., Ma, J.F., et al. (2007) The genetic basic and fine-mapping of a stable quantitative-trait loci for aluminium tolerance in rice. *Planta* 227, 255–262.
- Yamasaki, M., Tenaillon, M.I., Bi, I.V., Schroeder, S.G., Sanchez-Villeda, H., Doebley, J.F., et al. (2005) A large-scale screen for artificial selection in maize identifies candidate agronomic loci for domestication and crop improvement. *Plant Cell* 17, 2859–2872.
- Yan, L., Loukoianov, A., Tranquilli, G., Helguera, M., Fahima, T. and Dubcovsky, J. (2003) Positional cloning of the wheat vernalization gene *VRN1*. *Proc. Natl. Acad. Sci. USA* 100, 6263–6268.
- Yang, D.L., Jing, R.L., Chang, X.P. and Li, W. (2007a) Identification of quantitative trait loci and environmental interactions for accumulation and remobilization of water-soluble carbohydrates in wheat (*Triticum aestivum* L.) stems. *Genetics* 176, 571–584.
- Yang, D.L., Jing, R.L., Chang, X.P. and Li, W. (2007b) Quantitative trait loci mapping for chlorophyll fluorescence and associated traits in wheat (*Triticum aestivum*). *J. Integr. Plant Biol.* 49, 646–654.
- Yang, J., Sears, R.G., Gill, B.S. and Paulsen, G.M. (2002) Quantitative and molecular characterization of heat tolerance in hexaploid wheat. *Euphytica* 126, 275–282.
- Yang, L.T., Mickelson, S., See, D., Blake, T.K. and Fischer, A.M. (2004) Genetic analysis of the function of major leaf proteases in barley (*Hordeum vulgare* L.) nitrogen remobilization. *J. Exp. Bot.* 55, 2607–2616.
- Yang, Y., Ma, Y.Z., Xu, Z.S., Chen, X.M., He, Z.H., Yu, Z., et al. (2007a) Isolation and characterization of *Viviparous-1* genes in wheat cultivars with distinct ABA sensitivity and pre-harvest sprouting tolerance. *J. Exp. Bot.* 58, 2863–2871.
- Yang, Y., Zhao, X.L., Xia, L.Q., Chen, X.M., Xia, X.C., Yu, Z., et al. (2007b) Development and validation of a *Viviparous-1* STS marker for pre-harvest sprouting tolerance in Chinese wheats. *Theor. Appl. Genet.* 115, 971–980.
- Yau, S.K., Nachit, M. and Ryan, J. (1997) Variation in growth, development, and yield of durum wheat in response to high soil boron. II. Differences between genotypes. *Austr. J. Agric. Res.* 48, 951–957.
- Yin, X.Y., Stam, P., Kropff, M.J. and Schapendonk, A.H.C.M. (2003) Crop modeling, QTL mapping, and their complementary role in plant breeding. *Agron. J.* 95, 90–98.

- Yu, J.M., Holland, J.B., McMullen, M.D. and Buckler, E.S. (2008) Genetic design and statistical power of nested association mapping in maize. *Genetics* 178, 539–551.
- Yu, J.N., Huang, J., Wang, Z.N., Zhang, J.S. and Chen, S.Y. (2007) An Na^+/H^+ antiporter gene from wheat plays an important role in stress tolerance. *J. Biosci.* 32, 1153–1161.
- Zhang, H., Turner, N.C., Poole, M.L. and Simpson, N. (2006) Crop production in the high rainfall zones of southern Australia – potential, constraints and opportunities. *Austr. J. Agric. Res.* 46, 1035–1049.
- Zhou, L.L., Bai, G.H., Ma, H.X. and Carver, B.F. (2007) Quantitative trait loci for aluminum resistance in wheat. *Mol. Breed.* 19, 153–161.
- Zhu, B., Choi, D.W., Fenton, R. and Close, T.J. (2000) Expression of the barley dehydrin multigene family and the development of freezing tolerance. *Mol. Gen. Genet.* 264, 145–153.
- Zhu, J.K., Hasegawa, P. and Bressan, R. (1997) Molecular aspects of osmotic stress in plants. *Crit. Rev. Plant Sci.* 16, 253–277.
- Zhu, J.M., Alvarez, S., Marsh, E.L., LeNoble, M.E., Cho, I.J., Sivaguru, M., et al. (2007) Cell wall Proteome in the maize primary root elongation zone. II. Region-specific changes in water soluble and lightly ionically bound proteins under water deficit. *Plant Physiol.* 145, 1533–1548.

Chapter 19

Genomics of Biotic Interactions in the Triticeae

Roger P. Wise, Nick Lauter, Les Szabo, and Patrick Schweizer

Abstract In the area of Triticeae-pathogen interactions, highly parallel profiling of the transcriptome and proteome has provided entry points to examine host reaction to various pathogens and pests. In particular, the molecular mechanisms underlying gene-for-gene resistance and basal defense have been explored in the contrasting contexts of host vs. nonhost resistance and biotrophic vs. necrotrophic pathogenesis. Capitalizing on a rich history of genetics, molecular biology and plant pathology, recent studies in the Triticeae have provided new insights and characterized previously undescribed phenomena. The unique features of various pathosystems are increasingly leveraged by breakthroughs in genomic technologies, facilitating a community-wide approach to unifying themes of molecular plant-microbe interactions in the Triticeae.

19.1 Disease Epidemics and Current Threats

Plant diseases pose one of the greatest threats to agriculture in all corners of the world. In the Triticeae, major epidemics have occurred throughout history, affecting yield and quality of grain. Ergot of rye, which is caused by the ascomycete fungus, *Claviceps purpurea*, has possibly had more impact on the world than any other species of fungus. The earliest authenticated reports of its effects have been documented in Chinese writings around 1100 BC, and written accounts of wheat stem rust epidemics date back to ancient Greece and Rome.

R.P. Wise (✉)

Corn Insects and Crop Genetics Research, USDA-ARS; Department of Plant Pathology and Center for Plant Responses to Environmental Stresses, Iowa State University, Ames, Iowa 50011-1020
e-mail: rpwise@iastate.edu

In modern agricultural systems, the genetic uniformity required for mechanized production renders crops vulnerable to severe losses. Stem, leaf, and stripe rusts have been found in most areas of the world (Kolmer 2005). For example, stem rust, caused by the obligate fungal biotroph *Puccinia graminis* f. sp. *tritici* (*Pgt*), has been a serious problem wherever wheat and barley are grown (Leonard and Szabo 2005; Roelfs 1985), including a new threat by a novel race of wheat stem rust (*Pgt* TTKS) from East Africa (Stokstad 2007; Wanyera et al. 2006). This race, commonly called Ug99, is virulent on many current wheat varieties as well as advanced lines in most breeding programs (Bonman et al. 2007; Jin et al. 2007; Jin and Singh 2006). *Pgt* TTKS infects barley as well as wheat, thus, comprehensive molecular and genetic approaches will be necessary to provide a detailed understanding of the interactions of stem rust with its host plants, as well as a comparative resource to delineate conserved and unique mechanisms of resistance to Triticeae pathogens.

Other important pathogens and pests, such as *Fusarium graminearum* (Fusarium head blight or scab) (Parry et al. 1995), *Blumeria graminis* (powdery mildew) (Jørgensen 1994), *Pyrenophora teres* (net blotch) (Serenius et al. 2005), *Barley yellow dwarf virus* (BYDV) (D'Arcy and Burnett 1995), and Hessian Fly (Liu et al. 2007b) also have been serious deterrents to Triticeae grain production worldwide.

19.1.1 Plant Defenses Employed in Response to Biotic Stress

Historically, cereal crops have laid the foundation for numerous classical genetic studies in host-pathogen biology, resulting in many model biological systems. This is in part due to the longstanding and continuous threat posed by rusts and mildews. As a result, “gene-for-gene” and “non-host” interactions with wheat and barley have been primary targets for intricate biological studies.

In gene-for-gene systems, plants are protected from disease by specific recognition of diverse effectors presented by invading pathogens. This recognition capacity is encoded by plant-resistance (*R*) genes, which initiate signaling cascades through direct or indirect recognition of a cognate pathogen avirulence (*Avr*) gene product (Axtell and Staskawicz 2003; Mackey et al. 2003). The most prevalent class of plant *R* genes encode putative intracellular receptors containing highly conserved motifs including an N-terminal coiled coil (CC) or Toll/Interleukin-1 receptor-like (TIR) domain, a nucleotide binding site (NBS), and C-terminal, leucine rich repeats (LRR; Jones and Takemoto 2004; Jones and Dangl 2006). Specific residues within the LRR domains are hypervariable and targets for diversifying selection, which is a major factor in the determination of disease resistance specificity (Ellis et al. 2000, 2007; Michelmore and Meyers 1998; Mondragon-Palomino et al. 2002; Noel et al. 1999). In addition, genetic diversity also exists in the requirements for additional components of disease response pathways (Muskett and Parker 2003; Shirasu and

Schulze-Lefert 2003). Pathogen contact and recognition results in the induction of multiple components of the resistance response, including pathogenesis-related (*PR*) genes, the elicitation of systemic acquired resistance (SAR), and hypersensitive cell death (HR) (Hammond-Kosack and Jones 1996). These *R* genes (and their associated responses) are exploited by plant breeders to offset deleterious yield loss due to pathogen infection. However, mutations can occur within dynamic pathogen populations that alter the capacity of the host's *R* genes to recognize the invading pathogen.

Nonhost resistance functions at the plant and pathogen species level (Ellis 2006; Mellersh and Heath 2003; Thordal-Christensen 2003), and thus, has the potential for targeted deployment in Triticeae crops to control host pathogens. Two models of nonhost resistance are currently espoused. The first postulates the absence of adapted pathogen effectors, thereby leading to a non-compromised Pathogen Associated Molecular Patterns (PAMP)-triggered defense response, which is durable in nature. This type of resistance is also known as basal resistance or innate immunity in host systems. The second model postulates that nonhosts harbor multiple *R* genes that collectively recognize several to many pathogen-derived effector proteins (Jones and Dangl 2006). Such redundancy is thought to confer durable resistance because it is unlikely to be defeated on multiple fronts simultaneously, precluding positive selection for potentially virulent pathogen effector alleles that arise. In order to test these hypotheses in *Triticeae*, experimental approaches that overcome the obstacle of genetic separation of host- versus non-host plants, which usually belong to different species, will be essential.

19.1.2 Integrative Genomics Holds the Keys to Durable Resistance

Since the dawn of agriculture, the rate of crop improvement has been accelerating, making it possible for several billion people to eat without being involved in food production. In turn, trends toward globalization are heightening the epidemiological challenges facing agronomists and growers alike. Whereas *R*-genes could previously be used for crop protection until new variants arose in pathogen populations, migration of pre-existing virulent pathovars may become the more common source of defeat. This scenario emphasizes the general need to identify and utilize broader spectrum basal defense mechanisms that will confer durable resistance. A challenge herein is that these mechanisms often have only incremental effects on resistance when studied individually, making them much more difficult to identify experimentally. In order to create the germplasm that will thrive in these modern times, accelerated discovery of genic and biochemical targets and rapid application to crop improvement must be achieved. In this regard, transferring knowledge learned from model systems to agricultural applications lies in systems biology approaches that integrate complementary genomics experiments.

19.2 The Toolbox for Investigating Biotic Interactions

Genomic approaches commonly in use for investigating biotic interactions in the Triticeae generally fall into one of three categories: profiling of molecules to discover differential responses to treatments, integration of phenotypic, genetic and physical data to demonstrate consequences of an allelic difference, and systematically testing hypothetical functions of genes via stable and transient manipulations. Most investigations use a combination of these approaches, as they typically have complementary strengths. We review these in the context of the challenges posed and opportunities availed by interactions between pathogens and their Triticeae hosts.

19.2.1 Molecule Profiling Approaches

Molecule profiling approaches characterize biochemical and physiological differences that arise in response to treatment, genotype or both. Substantive variation manifests at all levels, including but not limited to differences in chromatin acetylation, DNA methylation, promoter function, splice variation, posttranscriptional antagonism, protein production and modification, sub-cellular targeting and finally, interaction with other required molecules. In both biological and technical terms, characterization of the transcriptome is by far the most tractable; it is the essential intermediate between the genome and the proteome and is readily characterized in its reverse transcribed state. Not surprisingly then, at least a plurality of current discoveries utilize transcriptomic experiments as entry points.

Profiling of mRNA in the Triticeae has been greatly facilitated by the development of Affymetrix GeneChips for both barley (Close et al. 2004) and wheat, facilitating acquisition of high quality data using platforms that gain utility as more experiments leverage them (Wise et al. 2007a). The Barley1 GeneChip assays mRNA levels for 21,439 unique genes and has been utilized extensively to analyze time-course expression profiles of hosts challenged by several pathogens (Wise et al. 2007b). While numerous discoveries have been made using this and other microarray platforms (see below), a limitation is that the expression levels for perhaps half of all barley genes are not queried. Although, expressed sequence tag libraries derived from infected materials were well represented in the unigene set to design the Barley1 GeneChip (Close et al. 2004), the application of next-generation deep sequencing technologies (e.g., 454 Life Science, Illumina/Solexa and ABISOLiD) to produce hundreds of millions of mRNA sequence tags will make more comprehensive microarrays possible within several years (Kling 2005; Margulies et al. 2005).

The Wheat GeneChip assays 55,052 transcripts, or about one third of all wheat genes, and is gaining momentum in the field of host-pathogen interactions (Coram et al. 2008; Liu et al. 2007b). However, proteomic profiling

by 2D electrophoresis has been more widely used thus far (Pós et al. 2005; Rampitsch et al. 2006; Wang et al. 2005; Zhou et al. 2005). This approach has a comprehensive potential in theory, but is limited by both sensitivity and electrophoretic resolution. It has the advantage of assaying gene products, rather than molecular intermediates, and can detect differences in phosphorylation states that can govern functionality if phosphoproteomic methods are employed (Rampitsch et al. 2006). A particular advantage of proteomic profiling in host-pathogen interaction studies is that the pathogen's proteome is also profiled. These studies will become increasingly valuable as plant and pathogen genomes are sequenced.

MicroRNAs have been shown to be an integral part of the defense response in *Arabidopsis* and loblolly pine, but characterization of their role in biotic stress response in Triticeae species has not yet been investigated (Lu et al. 2007; Navarro et al. 2006). However, recent microRNA profiling in wheat resulted in the identification of 23 new microRNA families, demonstrating the discovery power of deep sequencing approaches (Yao et al. 2007).

Success with profiling methods often depends on maximization of the signal-to-noise ratio. Single-cell transcriptomes have been profiled in barley to see what's different between a challenged and an unchallenged cell (Gjetting et al. 2004). In wheat, the intercellular secretome has been analyzed by proteomic methods to determine how defense responses at the point of contact are deployed (Pós et al. 2005). Even the method of RNA isolation can be optimized to better reflect the changing proteome in a transcriptome profiling experiment by hybridizing only cRNA associated with polysomes during translation (Skadsen and Jing 2008). Utilization of such biochemical properties will be required to identify the many types of molecular mechanisms underlying variation in defense response.

19.2.2 Integration of Phenotypic, Genetic and Physical-Map Data

Integration of phenotypic, genetic and physical-map associated sequence data to demonstrate consequences of allelic differences takes on divergent forms depending on whether the allelic differences in question are qualitative or quantitative. Map-based cloning (aka: positional cloning) integrates these data types to prove that a nucleotide sequence difference (typically qualitative) has a phenotypic consequence (see Chapter 12). This approach has been successfully used in the cloning of *R* genes responsible for triggering host resistance (Brueggeman et al. 2002; Halterman et al. 2001, 2003; Halterman and Wise 2004; Huang et al. 2003; Shen et al. 2003; Srichumpa et al. 2005; Wei et al. 1999, 2002; Yahiaoui et al. 2004).

Transcript-based cloning of genes using DNA microarrays has been used to discover important host-pathogen interactions as well. Expression values that are drastically lower in the mutant as compared to wild-type, such as those caused by deletion mutations, are the easiest to characterize (Mittra et al. 2004;

Zakhrabekova et al. 2002). This approach was used in barley to identify candidate genes encoding *Rpr1*, a gene required for *Rpg1*-mediated resistance to stem rust. A large deletion was identified via Barley1 GeneChip profiling of *rpr1* mutants derived from fast-neutron bombardment (Zhang et al. 2006). Three genes within the deletion cosegregated with the *rpr1*-mediated susceptible phenotype, narrowing the candidate gene list to facilitate characterization of *Rpg1*-specified resistance to stem rust.

Fast neutron mutants are well-suited for transcript-based cloning because the genetic lesions typically affect multiple genes and thus increase the likelihood of detecting expression knockouts (Alonso and Ecker 2006). However, since the deletions can vary in size from a single nucleotide to 30 kb, there is a potential of few to many genes that may need to be tested for co-segregation, and subsequent functional analysis performed.

When natural allelic differences are of interest, quantitative trait locus (QTL) mapping is typically the entry point. QTL mapping finds statistical associations between genotypes and phenotypes, allowing regions of the genome harboring allelic differences that cause variation in the phenotype to be identified (Mackay 2001). QTL mapping of infection type data has been widely used by plant pathologists to characterize the inheritance architecture of disease resistance and to identify chromosomal regions harboring regulators of defense (Horsley et al. 2005; Jiang et al. 2007; Leonova et al. 2007), but has not led to the cloning of genes other than those with qualitative effects. This is largely due to limitations in genetic resolution, which thwart efforts to singularly associate these causal differences with DNA sequence polymorphisms. Transcriptomic experiments tend to have a complementary deficiency; many DNA sequences are associated with differential response, but cause and effect relationships are not revealed. One of the most powerful approaches to elucidate consequences of natural allelic variation is to employ genetical genomics, a fusion of QTL mapping and high throughput genomic data collection techniques (Jansen and Nap 2001).

Expression QTL (eQTL) mapping is the most common genetical genomics approach. Transcript abundance of each single gene is treated as a quantitative trait and its regulation is genetically interrogated using QTL mapping approaches (Chen and Kendzierski 2007; Rockman and Kruglyak 2006). Since variation in gene expression has been shown to be a primary basis for the dynamic responses observed in plant-pathogen interactions, genetic interrogation of global gene expression during pathogen invasion is an appropriate way to identify and characterize defense gene networks (Hansen et al. 2008). By profiling gene expression in each member of a segregating population, it is possible to use linkage and network analyses to identify key regulators of gene expression for a particular condition (Jansen and Nap 2001; Rockman and Kruglyak 2006; Williams et al. 2007). eQTL mapping has been used in wheat to identify loci that regulate seed development (Jordan et al. 2007), and in barley to build gene networks associated with *R*-gene mediated and basal defense mechanisms involved in stem rust resistance (M. Moscou, N. Lauter, J. Rodriguez, G. Fuerst, B. Steffenson, Y. Jin, L. Szabo and R. Wise, unpublished results).

19.2.3 High-Throughput Functional Analysis

High-throughput functional capabilities for testing candidate genes are rapidly expanding. Functional validation of the candidate genes can be accomplished by genetic mutation, overexpression, or gene silencing (Caldwell et al. 2004; Douchkov et al. 2005; Hein et al. 2005; Scofield et al. 2005; Makandar et al. 2006). A number of resources needed for reverse genetic and functional analysis of candidate genes are now available for both plants and pathogens. Improvements in stable transformation (Fang et al. 2002; Hensel et al. 2008; Jacobsen et al. 2006; Janakiraman et al. 2002; Jones 2005) have benefited host-pathogen interaction studies just as they have other disciplines, but still remain a limitation. TILLING (Targeting Induced Local Lesions in Genomes, see Chapter 13) is also increasingly viable, but is often constrained by pathologically inappropriate genetic backgrounds, requiring introgression prior to functional characterization of alleles.

The most promising functional analysis breakthroughs for molecular plant pathologists are improvements in RNA interference (RNAi) techniques that alter gene expression in local cells or tissues that are the targets of pathogen invasion. Due to its sequence-homology dependent mode of action, double-stranded RNA interference (dsRNAi) is unique in its potential to overcome the problem of genetic redundancy, which is of critical importance for the polyploid genomes of Triticeae species (Chuang and Meyerowitz 2000; Schweizer et al. 2000; Tavernarakis et al. 2000).

Ten years ago, pioneering work by Bushnell and colleagues demonstrated the principle usefulness of a single-cell transient expression assay for powdery-mildew attacked barley (Nelson and Bushnell 1997). This assay was initially based on bombarded coleoptile tissue and subsequently developed for use in detached barley leaves (Nielsen et al. 1999; Schweizer et al. 1999b; Schweizer et al. 2000). This was possible because both the microprojectile-mediated transformation as well as powdery-mildew attack and potential development are cell-autonomous events taking place in single epidermal cells. The reliability of results obtained by using the transient assay has been verified by its ability to phenocopy allele introgression or gene mutation in barley and wheat, and by transgenic plants stably expressing genes of interest in leaf epidermis (Altpeter et al. 2005; Schultheiss et al. 2005; Schweizer et al. 2000).

The single-cell assay was further developed for transient-induced-gene-silencing (TIGS) by the expression of RNAi hairpin constructs (Douchkov et al. 2005). However, a limitation of the TIGS assay for barley-powdery mildew system has been the laborious phenotyping required to identify transformed cells and assign an infection severity rating based on haustorial features. Throughput is now significantly enhanced through use of microscope robotics and automated image analysis (Fig. 19.1) (Douchkov et al. 2005). Improvements in construct preparation using the GatewayTM technology have also accelerated this technique, such that 300 genes per person per

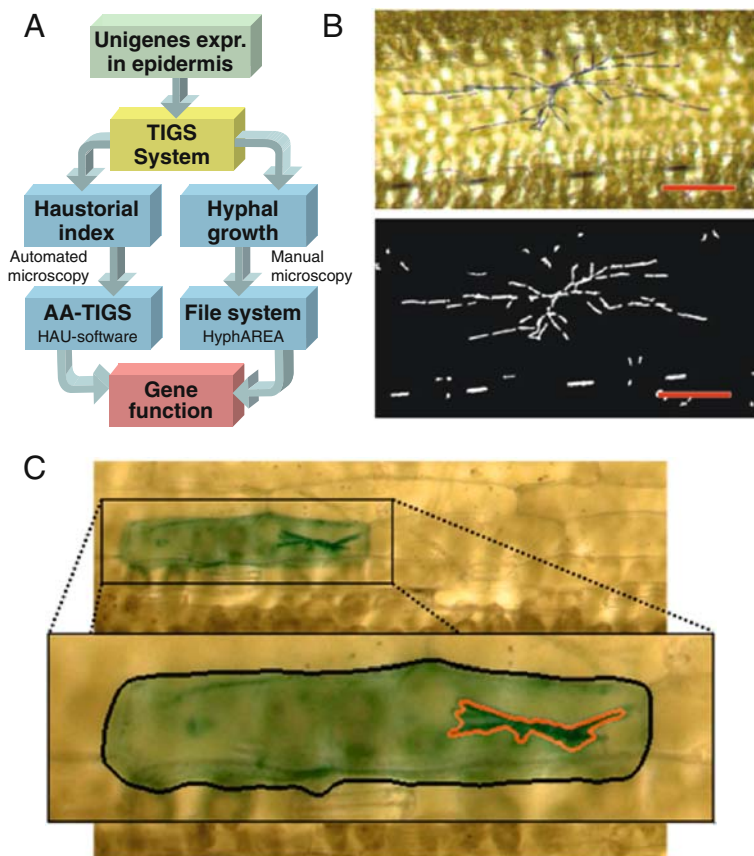


Fig. 19.1 High-throughput Transient-Induced-Gene-Silencing (TIGS) pipeline in barley. (A) Summary of approaches and tools available for quantitative phenomics in the barley-powdery mildew interaction: 1) Haustorial index is defined as the number of detected haustoria, divided by the number of observed GUS-positive, epidermal cells per bombardment. 2) Hyphal growth designates the increase of pixel numbers attributed over time to growing fungal colonies, as determined by the HyphAREA software tool. 3) File system designates the image files stored locally on a PC of data input into HyphAREA. 4) AA-TIGS and HAU-software is a Oracle database of microscopic images (input for the haustoria-recognition software) and quantitative microscopic data (output of the haustoria-recognition software) respectively. (B) Automated pattern recognition for quantitative assessment of fungal growth on the leaf surface. Upper panel, original image of a growing pustule, stained with Coomassie blue. Lower panel, final segmentation of hyphae for pixel quantification. Scale bar = 100 μ m. (C) Example of transformed epidermal cell expressing the *GUS* reporter gene prior to and after automated segmentation of cell (black bordering line) and haustorium (red bordering line). The figure was derived from a figure in Schweizer et al. (2006) *Hochdurchsatz-Phänomanalyse in Getreide*. GenomXPress 2/06: 16–19. GenomXPress is the non-indexed journal of the German genomics programs including GABI (See Color Insert)

month can be analyzed with the automated single-cell transient assay if two robotic microscopes are used. Since the initial reports, a considerable amount of functional information concerning defense-related barley and wheat genes has become available (Summarized in Table 19.1). In addition to the single-cell haustorium assay, an automated quantitative assessment of hyphal growth rate of powdery mildew has been developed, which can be used to study transgene effects manifested later in pathogenesis, permitting analysis of the slow mildewing phenomenon (Göllner et al. 2008; Seiffert and Schweizer 2005).

Table 19.1 Genes silenced or overexpressed using biolistic and virus-induced transient assays in Triticeae-fungal interactions (Updated from Panstruga 2004)

Category (Gene name) ^a	Description	OEx	RNAi ^b	Reference
Defense				
<i>TaPrx103 (WIR3)</i>	Peroxidase	X		(Schweizer et al. 1999b)
<i>TaWIR2</i>	Thaumatococcus	X		(Schweizer et al. 1999b)
<i>TaGluc</i>	β -1,3 glucanase	X		(Schweizer et al. 1999b)
<i>TaCH126</i>	Chitinase	X		(Schweizer et al. 1999b)
<i>TaGOX</i>	Glucose oxidase	X		(Schweizer et al. 1999b)
<i>HvPrx75; -85</i>	Peroxidase	X		(Kristensen et al. 2001)
<i>TaOXOX</i>	Oxalate oxidase	X		(Schweizer et al. 1999a)
<i>TaGLP4</i>	Germin-like	X		(Schweizer et al. 1999a)
<i>HvGER3; -4; -5</i>	Germin-like proteins	X	TIGS	(Zimmermann et al. 2006)
<i>HvADF3</i>	Actin-depol. factor	X		(Miklis et al. 2007)
Resistance signaling				
<i>HvMla</i> alleles	CC-NBS-LRR protein	X	TIGS/ VIGS	(Halterman et al. 2001; Halterman et al. 2003; Halterman and Wise 2004; Halterman and Wise 2006; Shen et al. 2003; Shen et al. 2007; Zhou et al. 2001)
<i>TaPm3</i> alleles	CC-NBS-LRR protein	X		(Srichumpa et al. 2005; Yahiaoui et al. 2006; Yahiaoui et al. 2004)
<i>HvSgt1; HvRar1</i>	SCF complex		TIGS/ VIGS	(Azevedo et al. 2002; Hein et al. 2005)
<i>TaSgt1; TaRar1</i>			VIGS	(Schofield et al. 2005)
<i>HvRacB</i>	Small GTP-binding protein	X	TIGS	(Schultheiss et al. 2003; Schultheiss et al. 2002)
<i>HvG (α)</i>	G protein subunit	X	TIGS	(Kim et al. 2002)

Table 19.1 (continued)

Category (Gene name) ^a	Description	OEx	RNAi ^b	Reference
Cell death				
<i>HvMlo</i>	7-Transmembrane protein	X	TIGS	(Kim et al. 2002; Schweizer et al. 2000; Shirasu et al. 1999a; Shirasu et al. 1999b)
<i>TaMlo</i>	7-Transmembrane protein	X		(Elliott et al. 2002)
<i>HvBI-1</i>	BAX inhibitor-like	X	TIGS	(Huckelhoven et al. 2003)
<i>HvCam-HvMlo</i>	Mlo-Calmodulin complex	X		(Kim et al. 2002)
<i>HvRBOH</i>	NADPH oxidase		TIGS	(Trujillo et al. 2006)
Transcription factor				
<i>HvNAC6</i>	NAC transcription factor	X	TIGS	(Jensen et al. 2007)
Avr_R_WRKY complex	Complex of <i>Bgh</i> AVR proteins with corresponding R proteins and WRKY transcription factors	X	VIGS	(Shen et al. 2007)
Vesicle traffic				
<i>HvSNAP34</i>	Syntaxin-interacting protein	X	TIGS	(Collins et al. 2003; Douchkov et al. 2005)
<i>HvUbi</i>	Ubiquitin	X	TIGS	(Dong et al. 2006)

^aThe prefix “*Ta*” or “*Hv*” specifies wheat or barley, respectively.

^bTIGS designates Transient Induced Gene Silencing. VIGS designates Virus Induced Gene Silencing.

More recently, virus induced gene silencing (VIGS) has emerged as a powerful reverse genetics tool for the functional analysis of gene candidates in both model and crop plant species. In monocots, *Brome mosaic virus* (BMV) has been utilized for functional genomics studies in rice and maize (Ding et al. 2006), whereas, *Barley stripe mosaic virus* (BSMV) has been developed for use in barley and wheat (Hein et al. 2005; Holzberg et al. 2002; Lacomme et al. 2003; Scofield et al. 2005). BSMV-VIGS has been used to silence genes involved in barley (Hein et al. 2005) and wheat (Zhou et al. 2007) defense by demonstrating resistance-breaking phenotypes in the host plants against powdery mildew and rust pathogens, respectively. As with bombardment, BSMV has also been developed as a novel vector for systemic transient overexpression expression, providing additional functional testing capabilities (Tai and Bragg 2007).

19.3 Triticeae-Fungal “Host” Interactions

Of the current threats, rusts and mildews are biotrophs, and *Fusarium* is a hemibiotroph and a necrotroph. Host-pathogen interaction studies have focused on biotrophs because of their significance, and also because the obligate nature of the interaction allows good control against accidental escape. As a result, several of the main themes in molecular plant-microbe interactions have been addressed using Triticeae-biotroph interactions. The fact that rice and maize are nonhosts for some of these diseases has allowed the further exploitation of these systems for the mining of biotic factors in disease defense.

Commonalities in host transcriptomic responses across diverse interactions between barley and the powdery mildew fungus, *Blumeria graminis* f. sp. *hordei* (*Bgh*), demonstrate that regardless of plant genotype, perception of general elicitors from pathogens leads to the induction of basal-defense responses (Caldo et al. 2004). These generalized defense mechanisms are thought to be targets for suppression by pathogen effector molecules deployed to establish infection (Alfano and Collmer 2004; Caldo et al. 2004; Espinosa and Alfano 2004). By extension, differences in the host gene expression patterns triggered by elicitors may provide insight into how pathogens respond and counter-attack early plant defense mechanisms (Caldo et al. 2006).

Transcriptomic investigations of *Mla*-mediated barley-powdery mildew interactions have also established a regulatory link between basal defense and *R*-gene mediated resistance, addressing one of the major questions outlined above (Caldo et al. 2004, 2006). A highly co-regulated cluster of >160 basal-defense related genes is significantly up-regulated in both incompatible and compatible interactions, coinciding with germination of *Bgh* conidia and formation of appressoria. Later, during establishment of the periahaustorial interface between penetrating *Bgh* and host epidermal cells, divergent expression of these transcripts occurs, where compatible interactions lead to lower expression compared to paired incompatible interactions. Many of these genes are associated with the basal defense response, which can be induced in a nonspecific manner by PAMPs (pathogen-associated-molecular-patterns) or MAMPs (microbe-associated-molecular-patterns) (Bent and Mackey 2007; Chisholm et al. 2006). The regulatory link between basal defense and *R*-gene mediated resistance has been confirmed by numerous subsequent studies (Kim et al. 2005; Navarro et al. 2004; Shen et al. 2007; Wise et al. 2007b). Significantly, the nuclear localization of MLA interactors indicates that two of these basal resistance regulators are a pair of transcription factors, *HvWRKY1* and *HvWRKY2* (Shen et al. 2007). Homologues of *HvWRKY1/2* in Arabidopsis, *AtWRKY18/40*, appear to be involved in a feedback repression system regulating deployment of basal defense (Shen et al. 2007).

Many outcomes are possible when thousands of powdery mildew conidia land on the same leaf. For example, one *Bgh* conidiospore may succeed during

the infection process and form a functional haustorium, whereas another may not. In order to investigate single-cell penetration attempts, Lyngkjaer and colleagues developed a system to micro-manipulate single barley epidermal cells (Gjetting et al. 2004) and examined differences in transcript accumulation following successful or unsuccessful penetration attempts. mRNA was captured on magnetic beads, purified and amplified for hybridization to the IPK (Gatersleben) barley PGRC1 10 k cDNA array (Sreenivasulu et al. 2001). Notably, up-regulation of sucrose synthase was exclusive to infected cells (Gjetting et al. 2007). In addition, two hexose transporters were up-regulated in both resistant and susceptible cells as well as genes associated with sucrose transport. Analogous findings have been observed in inoculated as compared to non-inoculated control samples from seedling leaves by Caldo and associates (2006); accession BB2 at <http://plexdb.org/>.

Bgh infects epidermal cells of barley, but also induces local acquired resistance in wheat, a nonhost (Bruggmann et al. 2005; Eichmann et al. 2006). To determine the extent to which epidermal penetration affects the disease response in mesophyll cells, incompatible responses of wheat were assayed in RNA samples enriched for epidermal or mesodermal leaf tissue. As expected, a large number of defense proteins were induced, but surprisingly, most had a greater fold change in the mesophyll than in the epidermis, indicating that local signaling occurs (Bruggmann et al. 2005).

Boddu and colleagues (Boddu et al. 2006), performed an extensive time-course study of barley cv. Morex spikes infected with *F. graminearum* isolate Butte 86 as compared to mock water control sampled at 24, 48, 72, 96, and 144 hai (Boddu et al. 2006; Accession no. BB9 at <http://plexdb.org/>). Interestingly, a large number of overlapping and conserved gene expression patterns were observed among these experiments and diverse host-pathogen interactions involving barley-*Bgh* (Caldo et al. 2004, 2006) and Arabidopsis response to its host powdery mildew, *Erysiphe orontii* (Wise et al. 2007b; Pathogenomics Integrated Microarray Database System: IMDS, <http://ausubellab.mgh.harvard.edu/imds/>, data downloaded November, 2006).

Trichothecene mycotoxin accumulation is associated with the shift from biotrophic to necrotrophic infection stages (Bushnell et al. 2003) and reduced virulence is correlated with low deoxynivalenol (DON) concentration in wheat (Mesterhazy et al. 1999). To understand the role of *F. graminearum* trichothecene mycotoxins in the barley host response, Boddu and associates (Boddu et al. 2007) investigated spikes of barley cv. Morex infected with trichothecene-producing wildtype strain (Z3639) and loss-of-function *tri5* trichothecene nonproducing mutant (GZT40). RNA was extracted at 48 and 96 hai for expression profiling (Boddu et al. 2007; Accession no. BB52 at <http://plexdb.org/>). This experiment revealed a unique set of 63 trichothecene-induced transcripts, in addition to accumulation of basal defense transcript observed by inoculation with the wild type strain of *F. graminearum*. These trichothecene-induced transcripts appear to control multiple cellular responses, including trichothecene detoxification and transport, ubiquitination, programmed

cell death, transcription, and secondary metabolism. Boddu et al. (2007) propose that trichothecene induces at least two classes of genes: (1) those involved in trichothecene detoxification, including UDPglucosyltransferases, ABC and MATE transporters, and (2) those encoding ubiquitination- and programmed-cell-death-related genes, including AAA-family ATPases, U-box domain protein, NF-X1 zinc finger protein, F-box protein, and a pirin protein. Thus, for the necrotrophic phase of *F. graminearum*, cell-death associated host gene expression likely enhances disease progression.

19.4 Triticeae-Fungal “Nonhost” Interactions

In the *Triticeae*, several wide interspecific introgressions of chromosome fragments confer strong and durable resistance in both wheat and barley (Mago et al. 2005; Oliver et al. 2005; Rubiales et al. 2001; Ruge et al. 2003; Tang et al. 1997). The basis for resistance in these cases is not clear; multiple *R* genes or major-effect “nonhost” defense genes are just two of the possibilities.

Progress has recently been obtained in barley attacked by inappropriate *formae speciales* (f. sp.) and species of the powdery mildew fungus *Blumeria graminis* and the rust fungus *Puccinia* sp., respectively. These obligate biotrophic rust and powdery mildew pathogens usually exhibit a high degree of host specificity, which means that they only infect one or a few closely-related hosts such as barley (*Hordeum* sp.) or wheat species (*Triticum* sp.). Therefore, a clear distinction of host *versus* nonhost interactions is possible allowing comparative or functional approaches. The focus on these pathosystems, together with the employment of multiple successive crosses, mutagenesis and gene-silencing approaches in barley opened up new possibilities to discover genes involved in nonhost resistance and to unravel the underlying mechanisms (Schweizer 2007).

In spring barley, the *mlo*-resistance gene has been widely used because it conferred race-nonspecific and durable resistance against *Bgh* for over 20 years. This provokes the question: Do *mlo*-mediated and nonhost resistance share common pathways that, in the case of *mlo*-mediated resistance, appear to result in strong basal resistance relieved from negative control by the MLO protein? There is ample evidence from physiological and transcript-profiling experiments in the barley-*Bgh* pathosystem that negative control of defense is indeed triggered or enhanced by successful fungal spores with MLO being one of the potential fungal effector (suppressor) targets (Lyngkjaer et al. 1997). The *PEN1* (for *penetration1*) and *Ror2* (for *required for mlo resistance 2*) genes have been found to encode functionally homologous tSNARE proteins in *Arabidopsis* and barley, respectively, that are predicted to mediate Golgi-vesicle fusion with target membranes (Collins et al. 2003). Although *Ror2* was originally identified in a mutant screen for breakdown of *mlo*-mediated host resistance to *Bgh*, it was recently found also to contribute to nonhost resistance, together with *Ror1*, another gene of unknown function required for *mlo*-mediated resistance

(Trujillo et al. 2004). The HvSNAP34 protein, which is similar to the yeast tSNARE protein SNAP25 (for synaptosome-associated protein of 25 kd) and predicted to interact with ROR2 in a ternary SNARE complex at target membranes, was found to be required for *mlo*-mediated host resistance against *Bgh* as well as for nonhost resistance to *B. graminis* f. sp. *tritici* (*Bgt*) (Douchkov et al. 2005). These results were obtained by transient-induced gene silencing (TIGS) in bombarded epidermal cells. The same dual functionality may be true for the *BI-1* gene encoding Bax-inhibitor 1, a putative negative regulator of cellular defense, and for the *ADF3* gene encoding the actin-depolymerizing factor 3 of barley because transient overexpression caused breakdown of nonhost- as well as *mlo*-mediated resistance (Eichmann et al. 2004; Huckelhoven et al. 2003; Miklis et al. 2007). Finally, *Mlo* overexpressing epidermal cells became susceptible to this inappropriate mildew suggesting that the cells use the same defense pathway that is under negative control by the MLO protein for basal host- and nonhost resistance (Elliott et al. 2002). In summary, mutant and reverse-genetic approaches in barley have produced results that suggest a functional overlap of *mlo*-mediated with nonhost resistance against powdery mildew.

In the interaction of barley with rust fungi, a better understanding of the genetic basis of nonhost resistance was achieved recently by accumulating alleles for nonhost susceptibility in a series of crosses, which resulted in two barley lines with essentially full susceptibility to inappropriate rust fungi (Atienza et al. 2004). This was possible because one of the parents of the initial cross, line L94, showed some marginal nonhost susceptibility to *P. triticina*. As a consequence, transgressive segregation of nonhost susceptibility was observed and exploited for consecutive crosses. Segregation analysis of progeny resulting from a series of crosses between “normal” nonhost resistant parents and one of the new, nonhost susceptible lines yielded a first insight into the genetic basis of the (non)host status of barley and revealed rather complex sets of mostly non-overlapping loci depending on the combination of parents (Jafary et al. 2006). Interestingly, map positions of a number of defense-related candidate genes in the progeny from these crosses were in significant association with nonhost-resistance QTLs. Therefore, it is tempting to speculate that not only in the interaction of barley with powdery mildew but also in barley-rust interactions, nonhost resistance is brought about by non-suppressed basal resistance that depends on a combination of defense-related alleles, which may vary from genotype to genotype and which reflects a considerable inherent degree of functional redundancy.

In contrast to barley, wheat might rather employ major *R*-genes with very durable effects for nonhost resistance to powdery mildews, as derived from *inter-formae-specialis* crosses of powdery mildew fungi. It was found that often one or a few segregating genes of the fungus were deciding upon host range, together with one or a few genes in the nonhost that co-segregated with known *R*-genes against *Bgt* such as *Pm10* (Matsumura and Tosa 1995; Tosa 1989). However, it remains a riddle why two as closely-related species as barley and wheat should differ in basal mechanisms of nonhost resistance against powdery

mildew fungi. Wheat would not be the only system in which *R*-gene mediated nonhost resistance has been observed because similar findings were reported for soybean (*Glycine max* L.) and *Arabidopsis* (Kobayashi et al. 1989; Staal et al. 2006). Since cultivated wheat has homoeologous A, B, and D genomes, an alternative explanation may be that knockout mutations specifying nonhost phenotypes have not yet been identified. In summary, the current data rather favor a model of nonhost resistance in barley that depends on a noncompromised, PAMP-triggered basal defense. This model is based on the assumption that fungal effectors (defense suppressors) released by inappropriate pathogens are largely ineffective against nonhosts. Future efforts on map-based cloning of nonhost QTLs and TILLING of candidate genes will provide a clearer picture of the genetic and molecular basis of nonhost resistance in Triticeae crops.

19.5 Triticeae Interactions with Insects, Viruses, Worms and Bacteria

Triticeae crops suffer significant damages from viruses, nematodes, insects and bacterial pathogens as well. *Barley yellow dwarf virus* (BYDV) and *Wheat streak mosaic virus* (WSMV) are potyviruses that cause significant economic losses, prompting major efforts to breed disease resistant varieties (Hakizimana et al. 2004; Sip et al. 2006). As is true for bacterial diseases, Triticeae species are not the best model hosts for genomics research on these interactions, so little is known about molecular genetic basis for the resistance mechanisms that breeders are using to protect small grain crops. The case is similar for Triticeae interaction with cereal cyst nematodes, as well as other parasitic worms that cause significant crop losses. Recently, several quantitative resistance loci have been reported and have been shown to confer meaningful levels of nematode resistance in wheat when pyramided together in elite lines (Barloy et al. 2007; Williams et al. 2006).

Wheat has emerged as a model host species for plant-insect interactions. Gene-for-gene interactions between Hessian fly (HF) and its wheat host, are specified by host *R* genes and insect *avr* genes (Harris et al. 2003). At least 32 *R*-gene loci conferring gene-for-gene resistance to HF have been identified in wheat, and in some cases finely mapped pursuant to positional cloning approaches (Kong et al. 2005, 2008; Nsarellah et al. 2003; Sardesai et al. 2005; Zhao et al. 2006). These cases of monogenic or oligogenic resistance have provided entry points into the molecular basis of plant defense against insects. Gene-for-gene interactions are infrequent among other plant-insect interactions, however. Hence, the question remains whether wheat also utilizes indirect defenses, for example, release of volatile chemicals to attract beneficial insectivores. Profiling of volatile chemicals from wheat plants infested by HF shows that these flies do not induce the production of volatiles that is typical of most insect herbivores, which may account for the observed failure of natural enemies to exert population control over HF in wheat crops (Tooker and De Moraes 2007).

Differential display was used to detect cDNAs that are upregulated in response to avirulent HF feeding on plants carrying the *Hf9* resistance gene (Williams et al. 2002). Several *Hessian fly-response* (*Hfr*) genes, *Hfr1*, *Hfr2* and *Hfr3*, have subsequently been shown to encode anti-nutritional lectins, proteins that bind various carbohydrates that affect insect feeding and digestion (Giovanini et al. 2007; Puthoff et al. 2005; Subramanyam et al. 2006). These appear to be relatively general defense response genes, as they can be at least partially induced by virulent and avirulent feeding by several different insects, mechanical wounding, and treatment with methyl jasmonate, salicylic acid and abscisic acid (Giovanini et al. 2007; Puthoff et al. 2005; Subramanyam et al. 2006).

An extensive mRNA profiling experiment involving three wheat cultivars and three HF biotypes was used to compare compatible and incompatible reactions as well as infested versus uninfested (Liu et al. 2007a). Notably, five lectin-domain containing genes, including *Hfr1*, were upregulated 5 to 15 fold in incompatible relative to compatible interactions. More importantly though, Liu and colleagues meticulously catalogued results from nearly 200 strongly differentially expressed genes that had putative functions associated with cell wall metabolism, antibiosis, phenylpropanoid biosynthesis, or nutrient metabolism and transport, among others. This list of gene functions is strikingly similar to the list produced by contrasting incompatible and compatible reactions to powdery mildew in barley, demonstrating the commonalities of *R*-gene mediated defense against a wide range of pathogens (Caldo et al. 2004, 2006).

19.6 Pathogen Genomics

Pathogen genomics has become tractable in the light of whole genome sequencing followed by bioinformatic interrogation to uncover genes encoding pathogen effectors, secreted proteins, and transcription factors that facilitate colonization of the host. In addition, transcriptome profiling has prompted a look into gene expression during pathogen attack. Below are updates on six of the major pathogens of the Triticeae.

19.6.1 *Fusarium graminearum* (*Fusarium Head Blight*)

Fusarium graminearum is the causal agent of head blight (scab) of wheat and barley. In the last decade *Fusarium* head blight (FHB) disease has rapidly re-emerged causing major epidemics in the North America and is becoming a threat globally (Goswami and Kistler 2004). *F. graminearum* (teleomorph *Gibberella zeae*) is a filamentous fungus, belonging to the phylum Ascomycota and order Hypocreales. The *F. graminearum* genome was whole-genome shotgun (WGS) sequenced (strain PH-1) and the resulting assembly totaled 36.45 Mb in 31 scaffolds (www.broad.mit.edu/annotation/genome/fusarium_group/MultiHome.html).

Genetic mapping of approximately 200 genome-sequence markers anchored 99.8% of the assembly to the genetic map (Gale et al. 2005). Current annotation of the genome contains 13,332 predicted genes. In contrast to other related fungi, *F. graminearum* genome contains few transposons and a low number of paralogues (Cuomo et al. 2007). The *F. graminearum* genome has greater number of genes for several categories, including transcription factors, hydrolytic enzymes and trans-membrane transporters when compared with *Magnaporthe grisea*, *Neurospora crassa* and *Aspergillus nigrus*. An extensive set of single-nucleotide polymorphisms (SNPs) were identified by comparing the assembly with a low coverage WGS from a second strain of *F. graminearum* (GZ3639). SNP densities varied across the genome sequence with the highest density in telomeric regions at the ends of the four chromosomes. In addition, three chromosomes had one or two regions with high SNP densities and correlated with regions of highest recombination. These regions of high SNP density have higher frequencies of genes specifically expressed during plant infection and lower frequencies of highly conserved genes.

Güldener and colleagues took advantage of the genome sequence of *F. graminearum* to design a whole-genome (18K) Affymetrix GeneChip (Güldener et al. 2006). To establish a baseline set of gene expression data, *F. graminearum* GeneChips were interrogated with RNA isolated from fungus grown in culture under three nutritional regimes: (a) complete medium (control), (b) minimal medium (Trail et al. 2003) without nitrogen, and (c) minimal medium without carbon, in addition to *in planta* growth in infected barley from the experiment performed by Boddu and associates (Boddu et al. 2006). Interestingly, 7,132 *Fusarium* probe sets were called present during the barley infection time course, even though the fraction of fungal transcripts in the total RNA from infected plants is quite low, notably during the early stages of infection. No enrichment was performed for fungal transcripts and the initial spore density used to inoculate plants was low. However, even at the earliest time point (24 hai), over 100 *Fusarium* sequences were detected.

Recently, WGS assemblies for two additional species of *Fusarium*, *F. oxysporum* and *F. verticillioides*, were completed (www.broad.mit.edu/annotation/genome/fusarium_group/MultiHome.html). Members of the *F. oxysporum* species complex cause vascular wilts of over 100 cultivated plant species including tomato, potato, sugarcane and cowpea, whereas, *F. verticillioides* is the causal agent of kernel and ear rot of maize. The genomes of *F. verticillioides* (41.7 Mb) and *F. oxysporum* (61.36 Mb) are larger than *F. graminearum* and are more complex containing higher levels of repetitive sequences and predicted genes.

19.6.2 *Puccinia graminis* (Stem Rust)

Puccinia graminis, the causal agent of stem rust, as a species has a relatively broad host range including more than 365 species of cereals and grasses

(Leonard and Szabo 2005). Historically, stem rust has been one of the most devastating diseases of wheat and barley. *P. graminis* has been divided into formae specialis based on host range which includes *P. graminis* f. sp. *tritici* (wheat and barley), *P. graminis* f. sp. *avenae* (oat) and *P. graminis* f. sp. *secalis* (rye). Like other rust fungi, *P. graminis* is functionally an obligate biotroph and therefore can only grow on living host tissue. The predominant asexual stage (uredinial) is dikaryotic (n+n) and occurs on its gramineous hosts, where sexual reproduction begins at the resting spore stage (telial) before culminating on the alternate host, barberry (*Berberis* spp.). *P. graminis* is a filamentous fungus, belonging to the phylum Basidiomycota and order Pucciniales.

The *P. graminis* f. sp. *tritici* (*Pgt*) genome was WGS sequenced (isolate CDL 75-36-700-3) using DNA from the dikaryotic urediniospores and the resulting assembly totaling 88.6 Mb (including 8% gaps) in 394 supercontigs (www.broad.mit.edu/annotation/genome/puccinia_graminis.2/). This represents the first sequencing of an obligate fungal plant pathogen. Forty-seven percent of the genome is composed of repetitive sequences, with transposons accounting for 12% of the repetitive sequences (C. Cuomo and L. Szabo, unpublished data). Current annotation predicts 20,567 genes covering 36% of the genome. Thirty-one percent of the predicted genes are supported by BLAST or *Pgt* EST data. The *Pgt* secretome is predicted to contain approximately 1,400 proteins, of which 73% are specific to the *Pgt* genome, which is likely to be reduced as additional genomes of rust fungi are sequenced. A partial genetic map has been developed, mapping eight avirulence loci on seven lineage groups (Zambino et al. 2000).

19.6.3 *Mycosphaerella graminicola* (*Septoria Tritici Blotch*)

Mycosphaerella graminicola (anamorph: *Septoria tritici*) is the causal agent of septoria tritici blotch, and is an agronomically important disease in most wheat growing regions worldwide. *M. graminicola* is a haploid, hemibiotrophic fungus with both filamentous and yeast-like growth phases. Early stage of infection is symptomless and intracellular (8–10 days) before a rapid collapse of mesophyll tissue occurs in a susceptible host. *M. graminicola* is an ascomycete in the class Dothideomycetes and order Capriodiales.

The *M. graminicola* genome was WGS sequenced (isolate IPO323) and the resulting assembly totaling 41.2 Mb in 129 scaffolds with roughly half of the genome contained in 6 scaffolds (<http://genome.jgipsf.org/Mycgr1/Mycgr1.home.html>). Annotation of this assembly includes a total of 11,396 predicted gene models. The current genetic map is composed of 26 linkage groups (Goodwin et al. 2007). The closely related *M. figiensis*, the causal agent of black leaf streak disease of bananas, has also been recently sequenced (<http://genome.jgi-psf.org/Mycf11/Mycf11.home.html>).

19.6.4 *Stagonospora nodorum* (*Stagonospora Nodorum Blotch*)

Stagonospora nodorum (teleomorph: *Phaeosphaeria nordorum*) is the causal agent of stagonospora nodorum blotch and is a major pathogen of wheat and barley (Solomon et al. 2006). The most damaging aspect of this disease is the infection of the head, leading to glume blotch. *S. nodorum* is an ascomycete in the class Dothideomycetes and order Pleosporales.

The *S. nordorum* genome was WGS sequenced (strain SN15) and the resulting assembly totaling 37.1 Mb in 107 scaffolds with more than 50% of the genome contained in the 13 largest (www.broad.mit.edu/annotation/genome/stagonospora_nodorum/Home.html). Approximately 4.5% of the genome is composed of repetitive sequences (Hane et al. 2007). Annotation of the assembled genome predicts at least 10,762 gene models covering 46% percent of the genome. Analysis of ESTs indicates that extracellular proteases, cellulases and xylanases predominate the fungal transcriptome in infected host tissue. The annotated *S. nodorum* genome encodes a large number of secreted proteins in which the majority show no significant homology with genes in current databases.

19.6.5 *Blumeria graminis* (*Powdery Mildew*)

Blumeria graminis, causal agent powdery mildew of cereals and grasses, has been subdivided into formae specialis based on host range. *B. graminis* f. sp. *hordei* and f. sp. *tritici* exclusively infects barley and wheat, respectively. Like all powdery mildews, *B. graminis* is an obligate biotrophic that, develop specialized feeding structure (haustoria). *Blumeria* epidemics result from a rapid (<5 day) asexual infection cycle, which culminates with the massive production of airborne conidia. The sexual cycle can be observed at the end of the host's growth season when environmental conditions become unfavorable. Outcrossing with compatible strains allows recombination and the generation of new variation on which selection pressure can act. *B. graminis* is an ascomycete in the class Leotiomycetes and order Erysiphales.

The genome of *B. graminis* f. sp. *hordei* (strain DH14) is currently being sequenced by BluGen, a BBSRC (UK)- sponsored international consortium (<http://www.blugen.org/>). In addition to sequence information on the genome, the website is also a repository for information, publications and images for powdery mildews. As of January 2009, the assembly spans 118 Mb. Initial analysis suggests that the size of the genome is approximately 120 Mb, however there is some uncertainty about the true value of this because of the exceptionally high level of repetitive DNA (60–70%) much of which is not found closely related sequenced ascomycetes (e.g. *Bortyis cinerea* and *Leptosphaeria maculans*). Preliminary data from high throughput sequencing of other powdery mildews appears to indicate that this might be a common feature amongst

Erysiphales. The BluGen project includes ESTs of conidia, conidia germinated on barley, infected epidermis (including haustoria), non-sporulating and sporulating epiphytic mycelia, and cleistothecia. These ESTs are a useful indication of the spectrum of genes expressed across different developmental stages and are being used to train the *Blumeria*-specific gene finding programs for the initial automated annotation.

Gene expression profiling using a custom *B. graminis* f. sp *hordei* cDNA microarray containing probes covering the *in planta* infection cycle in barley found patterns of coordinate expression among genes in defined metabolic pathways (Both et al. 2005). The authors were also able to assess the metabolic status of the fungus as it infects the host plant during asexual development. Genes encoding several glycolytic enzymes are significantly up-regulated as mature appressoria form. These genes are also seen in the infected epidermis, which contains fungal haustoria. These results correlate with host studies that show an up-regulation of sugar transport and utilization-related genes after infection with *B. graminis hordei* (Caldo, unpublished data; Gjetting et al. 2007) as well as *Arabidopsis* infected with *E. cichoracearum* (Fotopoulos et al. 2003).

19.6.6 Barley Yellow Dwarf Virus (BYDV)

Barley yellow dwarf virus (BYDV) and *Cereal yellow dwarf* (CYDV, formerly a strain of BYDV) are the most widespread and economically important viruses of wheat, barley and oats (D'Arcy and Burnett 1995; McKirdy et al. 2002). They occur wherever these crops are grown. The YDVs comprise a diverse mix of viruses (Miller et al. 2002) making resistance breeding difficult (Gray et al. 1994; Sharma et al. 1995). To begin to understand of the nature of YDVs around the world, the BYDV/CYDV Global Sequencing Project (PI's WA Miller, Iowa State University; SM Gray, Cornell University; J Anderson, Purdue University) commenced in 2004. Unlike projects that sequence multi-megabase genomes of bacterial or fungal pathogens, this small project focused on sequencing multiple, complete ~5,700 nt viral RNA genomes. Given that the viral genome is RNA with no poly(A) tail and that this virus occurs at low levels in plants, obtaining complete sequences was more time consuming and expensive per base than sequencing a DNA genome.

High quality sequences of complete genomes of over 40 YDV isolates from North America and Europe have been determined so far. These sequences revealed: (i) an isolate known as BYDV-RMV is neither BYDV nor CYDV but an entirely new virus; (ii) that a severe isolate, CYDV-RPS known previously to be only in Mexico is widespread throughout western North America and is an entirely different virus from CYDV-RPV; (iii) recombinants between CYDV-RPS and mild CYDV-RPV predominate in the Midwest; (iv) multiple recombination events occur in the coat protein genes suggesting a mechanism to generate virus isolates with new aphid vector specificities; (v) the most common

YDV, BYDV-PAV is remarkably invariant across North America; (vi) phylogenetic relationships of viruses based on structural (coat protein) genes is different from those based on the nonstructural (e.g. polymerase) genes because of the high recombination rate among genomes; (vii) there is a general trend toward less homology among related viruses at the 5' end of the genome compared to the middle of the genome.

19.7 Synthesis

The interactions between pathogens and plant hosts are a counter-balancing act; as the pathogen attempts to maximize nutrient siphoning and/or colonization, the host aims to restrict nutrient loss while minimizing the cost of defense. Plants initially activate basal defenses to limit penetration and colonization. However, pathogens are often capable of suppressing these innate defenses, requiring the plant to recognize them as threats in order to deploy stronger defenses.

These mechanisms evolved in natural populations, not in crops grown in monoculture. The defeat of recognition mechanisms by crop pathogens is now expected, requiring breeders to pyramid several *R* genes to ensure lasting protection against any single pathogen. Unfortunately, the threat of multiple pathogens in many growing regions makes it impractical to maintain optimal defense against each. A promising solution lies in the deployment of improved basal defense, which acts against a broad spectrum of pathogens and is not based on specific recognition, increasing its durability. But first, we must discover the molecular basis for more of these mechanisms across a wider spectrum of pathogens.

The state of the science for host-pathogen interactions in the Triticeae is encouraging. Well-developed experimental germplasm for targeted experiments in many emerging systems and the availability of large datasets from previous efforts present excellent opportunities for future analysis. The emergence of genetical genomic methods that attach cause to nucleic acid sequences, rather than simply to regions of the genome, is particularly encouraging. Likewise, advancements in high-throughput transient and virus-induced manipulations of gene expression make possible a wide variety of experiments that will be required to decipher intricacies of basal resistance, nonhost resistance, as well as race-specific resistance. These resources will promote the understanding of the complex nature of their interactions among Triticeae hosts and their pathogens, which will have long-term value for crop improvement.

Acknowledgments The authors thank Drs. Pietro Spanu and Allen Miller for their contributions of unpublished data on the *Bgh* and BYDV sequencing projects, respectively. Funding for this research was provided by USDA-ARS CRIS Project 3625-21000-049-00D (RW, NL), USDA Initiative for Future Agriculture and Food Systems (IFAFS) grant no. 2001-52100-11346 (RW), NSF Plant Genome Program # 0500461 "ISGA-Functional Genomics of Plant

Disease Defense Pathways” (RW), USDA-ARS CRIS Project 3640-21220-020-00D (LS), NSF Microbial Genome Sequencing grant EF-0412264 “Wheat Stem Rust Fungus Genome Sequencing Project” (LS), and German Ministry for Education and Research, Projects “BIC-GH-Bioinformatik Centrum Gatersleben-Halle” (PS), “GABI-nonhost” (PS), and BASF Plant Science Co. (PS).

This article is a joint contribution of the Corn Insects and Crop Genetics Research Unit, USDA-Agricultural Research Service, and The Iowa Agriculture and Home Economics Experiment Station. Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U.S. Department of Agriculture.

References

- Alfano, J.R. and Collmer, A. (2004) Type III secretion system effector proteins: double agents in bacterial disease and plant defense. *Annu. Rev. Phytopathol.* 42, 385–414.
- Alonso, J.M. and Ecker, J.R. (2006) Moving forward in reverse: genetic technologies to enable genome-wide phenomic screens in *Arabidopsis*. *Nat. Rev. Genet.* 7, 524–536.
- Altpeiter, F., Varshney, A., Abderhalden, O., Douchkov, D., Sautter, C., Kumlehn, J., Dudler, R. and Schweizer, P. (2005) Stable expression of a defense-related gene in wheat epidermis under transcriptional control of a novel promoter confers pathogen resistance. *Plant Mol. Biol.* 57, 271–283.
- Atienza, S.G., Jafary, H. and Niks, R.E. (2004) Accumulation of genes for susceptibility to rust fungi for which barley is nearly a nonhost results in two barley lines with extreme multiple susceptibility. *Planta* 220, 71–79.
- Axtell, M.J. and Staskawicz, B.J. (2003) Initiation of RPS2-specified disease resistance in *Arabidopsis* is coupled to the AvrRpt2-directed elimination of RIN4. *Cell* 112, 369–377.
- Azevedo, C., Sadanandom, A., Kitagawa, K., Freialdenhoven, A., Shirasu, K. and Schulze-Lefert, P. (2002) The RAR1 interactor SGT1, an essential component of *R* gene-triggered disease resistance. *Science* 295, 2073–2076.
- Barloy, D., Lemoine, J., Abelard, P., Tanguy, A., Rivoal, R. and Jahier, J. (2007) Marker-assisted pyramiding of two cereal cyst nematode resistance genes from *Aegilops variabilis* in wheat. *Mol. Breed.* 20, 31–40.
- Bent, A.F. and Mackey, D. (2007) Elicitors, effectors, and *R* genes: the new paradigm and a lifetime supply of questions. *Annu. Rev. Phytopathol.* 45, 399–436.
- Boddu, J., Cho, S., Kruger, W.M. and Muehlbauer, G.J. (2006) Transcriptome analysis of the barley-*Fusarium graminearum* interaction. *Mol. Plant Microbe Interact.* 19, 407–417.
- Boddu, J., Cho, S. and Muehlbauer, G.J. (2007) Transcriptome analysis of trichothecene-induced gene expression in barley. *Mol. Plant Microbe Interact.* 20, 1364–1375.
- Bonman, J.M., Bockelman, H.E., Jin, Y., Hijmans, R.J. and Gironella, A.I.N. (2007) Geographic distribution of stem rust resistance in wheat landraces. *Crop Sci.* 47, 1955–1963.
- Both, M., Csukai, M., Stumpf, M.P. and Spanu, P.D. (2005) Gene expression profiles of *Blumeria graminis* indicate dynamic changes to primary metabolism during development of an obligate biotrophic pathogen. *Plant Cell* 17, 2107–2122.
- Brueggeman, R., Rostoks, N., Kudrna, D., Kilian, A., Han, F., Chen, J., Druka, A., Steffenson, B. and Kleinbols, A. (2002) The barley stem rust-resistance gene *Rpg1* is a novel disease-resistance gene with homology to receptor kinases. *Proc. Natl. Acad. Sci. USA* 99, 9328–9333.
- Bruggmann, R., Abderhalden, O., Reymond, P. and Dudler, R. (2005) Analysis of epidermis- and mesophyll-specific transcript accumulation in powdery mildew-inoculated wheat leaves. *Plant Mol. Biol.* 58, 247–267.

- Bushnell, W.R., Hayen, B.E. and Pritsch, C. (2003) Histology and physiology of Fusarium head blight. In: K.J. Leonard and W.R. Bushnell (Eds.), *Fusarium Head Blight of Wheat And Barley*. American Phytopathological Society Press, St. Paul, MN, USA, pp. 44–83.
- Caldo, R.A., Nettleton, D., Peng, J. and Wise, R.P. (2006) Stage-specific suppression of basal defense discriminates barley plants containing fast- and delayed-acting *Mla* powdery mildew resistance alleles. *Mol. Plant Microbe Interact.* 19, 939–947.
- Caldo, R.A., Nettleton, D. and Wise, R.P. (2004) Interaction-dependent gene expression in *Mla*-specified response to barley powdery mildew. *Plant Cell* 16, 2514–2528.
- Caldwell, D.G., McCallum, N., Shaw, P., Muehlbauer, G.J., Marshall, D.F. and Waugh, R. (2004) A structured mutant population for forward and reverse genetics in barley (*Hordeum vulgare* L.). *Plant J.* 40, 143–150.
- Chen, M. and Kendziorski, C. (2007) A statistical framework for expression quantitative trait loci mapping. *Genetics* 177, 761–771.
- Chisholm, S.T., Coaker, G., Day, B. and Staskawicz, B.J. (2006) Host-microbe interactions: shaping the evolution of the plant immune response. *Cell* 124, 803–814.
- Chuang, C.F. and Meyerowitz, E.M. (2000) Specific and heritable genetic interference by double-stranded RNA in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA* 97, 4985–4990.
- Close, T.J., Wanamaker, S.I., Caldo, R.A., Turner, S.M., Ashlock, D.A., Dickerson, J.A., Wing, R.A., Muehlbauer, G.J., Kleinhofs, A. and Wise, R.P. (2004) A new resource for cereal genomics: 22 K barley GeneChip comes of age. *Plant Physiol.* 134, 960–968.
- Collins, N.C., Thordal-Christensen, H., Lipka, V., Bau, S., Kombrink, E., Qiu, J.L., Huckelhoven, R., Stein, M., Freialdenhoven, A., Somerville, S.C. and Schulze-Lefert, P. (2003) SNARE-protein-mediated disease resistance at the plant cell wall. *Nature* 425, 973–977.
- Coram, T.E., Wang, M. and Chen, X. (2008) Transcriptome analysis of the wheat-*Puccinia striiformis* f. sp. *tritici* interaction. *Mol. Plant Pathol.* 9, 157–169.
- Cuomo, C.A., Guldener, U., Xu, J.-R., Trail, F., Turgeon, B.G., Pietro, A.D., Walton, J., D., Ma, L.-J., Baker, S.E., Rep, M., Adam, G., Antoniw, J., Baldwin, T., Calvo, S., Chang, Y.-L., DeCaprio, D., Gale, L.R., Gnerre, S., Goswami, R.S., Hammond-Kosack, K., Harris, L.J., Hilburn, K., Kennell, J.C., Kroken, S., Magnuson, J.K., Mannhaupt, G., mauceli, E., Mewes, H.-W., Mitterbauer, G., Munsterkotter, M., Nelson, D., O'Donnell, K., Oueller, T., Qi, W., Quesneville, H., Roncero, M.I.G., Seong, K.-Y., Tetko, I.V., Urban, M., Wallwijk, C., Ward, T.J., Yao, J., Birren, B.W. and Kistler, H.C. (2007) The *Fusarium graminearum* genome reveals a link between localized polymorphism and pathogen specialization. *Science* 317, 1400–1402.
- D'Arcy, C.J. and Burnett, P.A. (Eds.). (1995) *Barley Yellow Dwarf: 40 Years of Progress*. APS Press, St. Paul.
- Ding, X.S., Schneider, W.L., Chaluvadi, S.R., Mian, M.A. and Nelson, R.S. (2006) Characterization of a *Brome mosaic virus* strain and its use as a vector for gene silencing in monocotyledonous hosts. *Mol. Plant Microbe Interact.* 19, 1229–1239.
- Dong, W.B., Nowara, D. and Schweizer, P. (2006) Protein polyubiquitination plays a role in basal host resistance of barley. *Plant Cell* 18, 3321–3331.
- Douchkov, D., Nowara, D., Zierold, U. and Schweizer, P. (2005) A high-throughput gene-silencing system for the functional assessment of defense-related genes in barley epidermal cells. *Mol. Plant Microbe Interact.* 18, 755–761.
- Eichmann, R., Biemelt, S., Schafer, P., Scholz, U., Jansen, C., Felk, A., Schafer, W., Langen, G., Sonnewald, U., Kogel, K.-H. and Huckelhoven, R. (2006) Macroarray expression analysis of barley susceptibility and nonhost resistance to *Blumeria graminis*. *J. Plant Physiol.* 163, 657–670.
- Eichmann, R., Schultheiss, H., Kogel, K.H. and Huckelhoven, R. (2004) The barley apoptosis suppressor homologue BAX inhibitor-1 compromises nonhost penetration resistance of barley to the inappropriate pathogen *Blumeria graminis* f. sp. *tritici*. *Mol. Plant Microbe Interact.* 17, 484–490.

- Elliott, C., Zhou, F.S., Spielmeyer, W., Panstruga, R. and Schulze-Lefert, P. (2002) Functional conservation of wheat and rice *Mlo* orthologs in defense modulation to the powdery mildew fungus. *Mol. Plant Microbe Interact.* 15, 1069–1077.
- Ellis, J. (2006) Insights into nonhost disease resistance: can they assist disease control in agriculture? *Plant Cell* 18, 523–528.
- Ellis, J., Dodds, P. and Pryor, T. (2000) Structure, function and evolution of plant disease resistance genes. *Curr. Opin. Plant Biol.* 3, 278–284.
- Ellis, J.G., Dodds, P.N. and Lawrence, G.J. (2007) Flax rust resistance gene specificity is based on direct resistance-avirulence protein interactions. *Annu. Rev. Phytopathol.* 45, 289–306.
- Espinosa, A. and Alfano, J.R. (2004) Disabling surveillance: bacterial type III secretion system effectors that suppress innate immunity. *Cell. Microbiol.* 6, 1027–1040.
- Fang, Y.-D., Akula, C. and Altpeter, F. (2002) *Agrobacterium*-mediated barley (*Hordeum vulgare* L.) transformation using green fluorescent protein as a visual marker and sequence analysis of the T-DNA::barley genomic DNA junctions. *J. Plant Physiol.* 159, 1131–1138.
- Fotopoulos, V., Gilbert, M.J., Pittman, J.K., Marvier, A.C., Buchanan, A.J., Sauer, N., Hall, J.L. and Williams, L.E. (2003) The monosaccharide transporter gene, *AtSTP4*, and the cell-wall invertase, *Atbetafruct1*, are induced in *Arabidopsis* during infection with the fungal biotroph *Erysiphe cichoracearum*. *Plant Physiol.* 132, 821–829.
- Gale, L.R., Bryant, J.D., Calvo, S., Giese, H., Katan, T., O'Donnell, K., Suga, H., Taga, M., Usgaard, T.R., Ward, T.J. and Kistler, H.C. (2005) Chromosome complement of the fungal plant pathogen *Fusarium graminearum* based on genetic and physical mapping and cytological observations. *Genetics* 171, 985–1001.
- Giovanini, M.P., Saltzmann, K.D., Puthoff, D.P., Gonzalo, M., Ohm, H.W. and Williams, C.E. (2007) A novel wheat gene encoding a putative chitin-binding lectin is associated with resistance against Hessian fly. *Mol. Plant Pathol.* 8, 69–82.
- Gjetting, T., Carver, T.L., Skot, L. and Lyngkjær, M.F. (2004) Differential gene expression in individual papilla-resistant and powdery mildew-infected barley epidermal cells. *Mol. Plant Microbe Interact.* 17, 729–738.
- Gjetting, T., Hagedorn, P.H., Schweizer, P., Thordal-Christensen, H., Carver, T.L.W. and Lyngkjær, M.F. (2007) Single-cell transcript profiling of barley attacked by the powdery mildew fungus. *Mol. Plant Microbe Interact.* 20, 235–246.
- Göllner, K., Schweizer, P., Bai, Y. and Panstruga, R. (2008) Natural genetic resources of *Arabidopsis thaliana* reveal a high prevalence and unexpected plasticity of *RPW8*-mediated powdery-mildew resistance. *New Phytol.* 177, 725–742.
- Goodwin, S.B., van der Lee, T.A.J., Cavaletto, J.R., Hekkert, B.t.L., Crane, C.F. and Kerma, G.H.J. (2007) Identification and genetic mapping of highly polymorphic microsatellite loci from an EST database of the septoria tritici blotch pathogen *Mycosphaerella graminicola*. *Fungal Genet. Biol.* 44, 398–414.
- Goswami, R.S. and Kistler, H.C. (2004) Heading for disaster: *Fusarium graminearum* on cereal crops. *Mol. Plant Pathol.* 5, 515–525.
- Gray, S.M., Smith, D. and Sorrells, M. (1994) Reduction of disease incidence in small field plots by isolate-specific resistance to barley yellow dwarf virus. *Phytopathology* 84, 713–718.
- Göldner, U., Seong, K.-Y., Boddu, J., Cho, S., Trail, F., Xu, J.-R., Adam, G., Mewes, H.-W., Muehlbauer, G.J. and Kistler, H.C. (2006) Development of a *Fusarium graminearum* Affymetrix GeneChip for profiling fungal gene expression *in vitro* and *in planta*. *Fungal Genet. Biol.* 43, 316–325.
- Hakizimana, F., Ibrahim, A.M.H., Langham, M.A.C., Haley, S.D. and Rudd, J.C. (2004) Diallel analysis of wheat streak mosaic virus resistance in winter wheat. *Crop Sci.* 44, 89–92.
- Halterman, D., Zhou, F., Wei, F., Wise, R.P. and Schulze-Lefert, P. (2001) The MLA6 coiled-coil, NBS-LRR protein confers *AvrMla6*-dependent resistance specificity to *Blumeria graminis* f. sp. *hordei* in barley and wheat. *Plant J.* 25, 335–348.

- Halterman, D.A., Wei, F. and Wise, R.P. (2003) Powdery mildew-Induced *Mla* mRNAs are alternatively spliced and contain multiple upstream open reading frames. *Plant Physiol.* 131, 558–567.
- Halterman, D.A. and Wise, R.P. (2004) A single-amino acid substitution in the sixth leucine-rich repeat of barley *MLA6* and *MLA13* alleviates dependence on *RAR1* for disease resistance signaling. *Plant J.* 38, 215–226.
- Halterman, D.A. and Wise, R.P. (2006) Upstream open reading frames of the barley *Mla13* powdery mildew resistance gene function co-operatively to down-regulate translation. *Mol. Plant Pathol.* 7, 167–176.
- Hammond-Kosack, K.E. and Jones, J.D. (1996) Resistance gene-dependent plant defense responses. *Plant Cell* 8, 1773–1791.
- Hane, J.K., Lowe, R.G.T., Solomon, P.S., Tan, K.-C., Schoch, C.L., Spatafora, J.W., Crous, P.W., Kodira, C., Birren, B.W., Galagan, J.E., Torriani, S.F.F., McDonald, B. A. and Oliver, R.P. (2007) Dothideomycete-Plant interactions illuminated by genome sequencing and EST analysis of the wheat pathogen *Stagonospora nodorum*. *Plant Cell* 19, 3347–3368.
- Hansen, B.G., Halkier, B.A. and Kliebenstein, D.J. (2008) Identifying the molecular basis of QTLs: eQTLs add a new dimension. *Trends Plant Sci.* 13, 72–77.
- Harris, M.O., Stuart, J.J., Mohan, M., Nair, S., Lamb, R.J. and Rohfritsch, O. (2003) Grasses and gall midges: plant defense and insect adaptation. *Annu. Rev. Entomol.* 48, 549–577.
- Hein, I., Barciszewska-Pacak, M., Hrubikova, K., Williamson, S., Dinesen, M., Soenderby, I. E., Sundar, S., Jarmolowski, A., Shirasu, K. and Lacomme, C. (2005) Virus-induced gene silencing-based functional characterization of genes associated with powdery mildew resistance in barley. *Plant Physiol.* 138, 2155–2164.
- Hensel, G., Valkov, V., Middlefell-Williams, J. and Kumlehn, J. (2008) Efficient generation of transgenic barley: the way forward to modulate plant-microbe interactions. *J. Plant Physiol.* 165, 71–82.
- Holzberg, S., Brosio, P., Gross, C. and Pogue, G.P. (2002) Barley stripe mosaic virus-induced gene silencing in a monocot plant. *Plant J.* 30, 315–327.
- Horsley, R.D., Schmierer, D., Maier, C., Kudrna, D., Urrea, C.A., Steffenson, B.J., Schwarz, P.B., Franckowiak, J.D., Green, M.J., Zhang, B. and Kleinhofs, A. (2005) Identification of QTLs associated with *Fusarium* head blight resistance in barley accession CIho 4196. *Crop Sci.* 46, 145–156.
- Huang, L., Brooks, S.A., Li, W., Fellers, J.P., Trick, H.N. and Gill, B.S. (2003) Mapbased cloning of leaf rust resistance gene *Lr21* from the large and polyploid genome of bread wheat. *Genetics* 164, 655–664.
- Huckelhoven, R., Dechert, C. and Kogel, K.H. (2003) Overexpression of barley BAX inhibitor 1 induces breakdown of *mlo*-mediated penetration resistance to *Blumeria graminis*. *Proc. Natl. Acad. Sci. USA* 100, 5555–5560.
- Jacobsen, J., Venables, I., Wang, M.-B., Matthews, P., Ayliffe, M. and Gubler, F. (2006) Barley (*Hordeum vulgare* L.). In: Wang, K. (Ed.), *Agrobacterium Protocols*. Humana press, New Jersey, pp. 171–184.
- Jafary, H., Szabo, L.J. and Niks, R.E. (2006) Innate nonhost immunity in barley to different heterologous rust fungi is controlled by sets of resistance genes with different and overlapping specificities. *Mol. Plant Microbe Interact.* 19, 1270–1279.
- Janakiraman, V., Steinau, M., McCoy, S. and Trick, H. (2002) Recent advances in wheat transformation. *In Vitro Cell. Dev. Biol. Plant* 38, 404–414.
- Jansen, R.C. and Nap, J.P. (2001) Genetical genomics: the added value from segregation. *Trends Genet.* 17, 388–391.
- Jensen, M.K., Rung, J.H., Gregersen, P.L., Gjetting, T., Fuglsang, A.T., Hansen, M., Joehnk, N., Lyngkjaer, M.F. and Collinge, D.B. (2007) The HvNAC6 transcription factor: a positive regulator of penetration resistance in barley and Arabidopsis. *Plant Mol. Biol.* 65, 137–150.

- Jiang, G.-L., Dong, Y., Shi, J. and Ward, R. (2007) QTL analysis of resistance to Fusarium head blight in the novel wheat germplasm CJ 9306. II. Resistance to deoxynivalenol accumulation and grain yield loss. *Theor. Appl. Genet.* 115, 1043–1052.
- Jin, Y., Pretorius, Z.A. and Singh, R.P. (2007) New virulence within race TTKS (Ug99) of the stem rust pathogen and effective resistance genes. *Phytopathology* 97, S137.
- Jin, Y. and Singh, R.P. (2006) Resistance in U.S. wheat to recent eastern African isolates of *Puccinia graminis* f. sp. *tritici* with virulence to resistance gene *Sr31*. *Plant Dis.* 90, 476–480.
- Jones, D.A. and Takemoto, D. (2004) Plant innate immunity – direct and indirect recognition of general and specific pathogen-associated molecules. *Curr. Opin. Immunol.* 16, 48–62.
- Jones, H.D. (2005) Wheat transformation: current technology and applications to grain development and composition. *J. Cereal Sci.* 41, 137–147.
- Jones, J.D.G. and Dangl, J.L. (2006) The plant immune system. *Nature* 444, 323–329.
- Jordan, M.C., Somers, D.J. and Banks, T.W. (2007) Identifying regions of the wheat genome controlling seed development by mapping expression quantitative trait loci. *Plant Biotechnol. J.* 5, 442–453.
- Jørgensen, J.H. (1994) Genetics of powdery mildew resistance in barley. *Crit. Rev. Plant Sci.* 13, 97–119.
- Kim, M.C., Panstruga, R., Elliott, C., Muller, J., Devoto, A., Yoon, H.W., Park, H.C., Cho, M.J. and Schulze-Lefert, P. (2002) Calmodulin interacts with MLO protein to regulate defence against mildew in barley. *Nature* 416, 447–451.
- Kim, M.G., da Cunha, L., McFall, A.J., Belkadir, Y., DebRoy, S., Dangl, J.L. and Mackey, D. (2005) Two *Pseudomonas syringae* Type III effectors inhibit RIN4-regulated basal defense in *Arabidopsis*. *Cell* 121, 749–759.
- Kling, J. (2005) The search for a sequencing thoroughbred. *Nat. Biotechnol.* 23, 1333–1335.
- Kobayashi, D.Y., Tamaki, S.J. and Keen, N.T. (1989) Cloned avirulence genes from the tomato pathogen *Pseudomonas syringae* pathovar tomato confer cultivar specificity on soybean. *Proc. Natl. Acad. Sci. USA* 86, 157–161.
- Kolmer, J.A. (2005) Tracking wheat rust on a continental scale. *Curr. Opin. Plant Biol.* 8, 441–449.
- Kong, L., Cambron, S. and Ohm, H. (2008) Hessian fly resistance genes *H16* and *H17* are mapped to a resistance gene cluster in the distal region of chromosome 1AS in wheat. *Mol. Breed.* 21, 183–194.
- Kong, L., Ohm, H.W., Cambron, S.E. and Williams, C.E. (2005) Molecular mapping determines that Hessian fly resistance gene *H9* is located on chromosome 1A of wheat. *Plant Breed.* 124, 525–531.
- Kristensen, B.K., Ammitzboll, H., Rasmussen, S.K. and Nielsen, K.A. (2001) Transient expression of a vacuolar peroxidase increases susceptibility of epidermal barley cells to powdery mildew. *Mol. Plant Pathol.* 2, 311–317.
- Lacomme, C., Hrubikova, K. and Hein, I. (2003) Enhancement of virus-induced gene silencing through viral-based production of inverted-repeats. *Plant J.* 34, 543–553.
- Leonard, K.J. and Szabo, L.J. (2005) Stem rust of small grains and grasses caused by *Puccinia graminis*. *Mol. Plant Pathol.* 6, 99–111.
- Leonova, I., Laikova, L., Popova, O., Unger, O., Börner, A. and Röder, M. (2007) Detection of quantitative trait loci for leaf rust resistance in wheat—*T. timopheevii*/*T. tauschii* introgression lines. *Euphytica* 155, 79–86.
- Liu, X., Bai, J., Huang, L., Zhu, L., Liu, X., Weng, N., Reese, J., Harris, M., Stuart, J. and Chen, M.-S. (2007a) Gene expression of different wheat genotypes during attack by virulent and avirulent Hessian fly (*Mayetiola destructor*) larvae. *J. Chem. Ecol.* 33, 2171–2194.
- Liu, X., Jianfa, B., Huang, L., Zhu, L., Liu, X., Weng, N., Reese, J.C., Harris, M., Stuart, J.J. and Chen, M.-S. (2007b) Gene expression of different wheat genotypes during attack by virulent and avirulent hessian fly (*Mayetiola destructor*) larvae. *J. Chem. Ecol.* 33, 2171–2194.

- Lu, S., Sun, Y.-H., Amerson, H. and Chiang, V.L. (2007) MicroRNAs in loblolly pine (*Pinus taeda* L.) and their association with fusiform rust gall development. *Plant J.* 51, 1077–1098.
- Lyngkjaer, M.F., Carver, T.L.W. and Zeyen, R.J. (1997) Suppression of resistance to *Erysiphe graminis* f. sp. *hordei* conferred by the *mlo5* barley powdery mildew resistance gene. *Physiol. Mol. Plant Pathol.* 50, 17–36.
- Mackay, T.F.C. (2001) The genetic architecture of quantitative traits. *Annu. Rev. Genet.* 35, 303–339.
- Mackey, D., Belkhadir, Y., Alonso, J.M., Ecker, J.R. and Dangl, J.L. (2003) Arabidopsis RIN4 is a target of the type III virulence effector AvrRpt2 and modulates RPS2-mediated resistance. *Cell* 112, 379–389.
- Mago, R., Miah, H., Lawrence, G.J., Wellings, C.R., Spielmeier, W., Bariana, H.S., McIntosh, R.A., Pryor, A.J. and Ellis, J.G. (2005) High-resolution mapping and mutation analysis separate the rust resistance genes *Sr31*, *Lr26* and *Yr9* on the short arm of rye chromosome 1. *Theor. Appl. Genet.* 112, 41–50.
- Makandar, R., Essig, J.S., Schapaugh, M.A., Trick, H.N. and Shah, J. (2006) Genetically engineered resistance to Fusarium head blight in wheat by expression of Arabidopsis NPR1. *Mol. Plant Microbe Interact.* 19, 123–129.
- Margulies, M., Egholm, M., Altman, W.E., Attiya, S., Bader, J.S., Bemben, L.A., Berka, J., Braverman, M.S., Chen, Y.J., Chen, Z., Dewell, S.B., Du, L., Fierro, J.M., Gomes, X.V., Godwin, B.C., He, W., Helgesen, S., Ho, C.H., Irzyk, G.P., Jando, S.C., Alenquer, M.L., Jarvie, T.P., Jirage, K.B., Kim, J.B., Knight, J.R., Lanza, J.R., Leamon, J.H., Lefkowitz, S.M., Lei, M., Li, J., Lohman, K.L., Lu, H., Makhijani, V.B., McDade, K.E., McKenna, M.P., Myers, E.W., Nickerson, E., Nobile, J.R., Plant, R., Puc, B.P., Ronan, M.T., Roth, G.T., Sarkis, G.J., Simons, J.F., Simpson, J.W., Srinivasan, M., Tartaro, K.R., Tomasz, A., Vogt, K.A., Volkmer, G.A., Wang, S.H., Wang, Y., Weiner, M.P., Yu, P., Begley, R.F. and Rothberg, J.M. (2005) Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437, 376–380.
- Matsumura, K. and Tosa, Y. (1995) The rye mildew fungus carries avirulence genes corresponding to wheat genes for resistance to races of the wheat mildew fungus. *Phytopathology* 85, 753–756.
- McKirdy, S.J., Jones, R.A.C. and Nutter, F.W. (2002) Quantification of yield losses caused by *Barley yellow dwarf virus* in wheat and oats. *Phytopathology* 86, 769–773.
- Mellersh, D. and Heath, M. (2003) An investigation into the involvement of defense signaling pathways in components of the nonhost resistance of *Arabidopsis thaliana* to rust fungi also reveals a model system for studying rust fungal compatibility. *Mol. Plant Microbe Interact.* 16, 398–404.
- Mesterhazy, A., Bartok, T., Mirocha, C.G. and Komoroczy, R. (1999) Nature of wheat resistance to Fusarium head blight and the role of deoxynivalenol for breeding. *Plant Breed.* 118, 97–110.
- Michelmore, R.W. and Meyers, B.C. (1998) Clusters of resistance genes in plants evolve by divergent selection and a birth-and-death process. *Genome Res.* 8, 1113–1130.
- Miklis, M., Consonni, C., Bhat, R.A., Lipka, V., Schulze-Lefert, P. and Panstruga, R. (2007) Barley MLO modulates actin-dependent and actin-independent antifungal defense pathways at the cell periphery. *Plant Physiol.* 144, 1132–1143.
- Miller, W.A., Liu, S. and Beckett, R. (2002) Barley yellow dwarf virus: *Luteoviridae* or *Tombusviridae*? *Mol. Plant Pathol.* 3, 177–183.
- Mitra, R.M., Gleason, C.A., Edwards, A., Hadfield, J., Downie, J.A., Oldroyd, G.E. and Long, S.R. (2004) A Ca²⁺/calmodulin-dependent protein kinase required for symbiotic nodule development: gene identification by transcript-based cloning. *Proc. Natl. Acad. Sci. USA* 101, 4701–4705.
- Mondragon-Palomino, M., Meyers, B.C., Michelmore, R.W. and Gaut, B.S. (2002) Patterns of positive selection in the complete NBS-LRR gene family of *Arabidopsis thaliana*. *Genome Res.* 12, 1305–1315.

- Muskett, P. and Parker, J. (2003) Role of SGT1 in the regulation of plant R gene signalling. *Microbes Infect.* 5, 969–976.
- Navarro, L., Dunoyer, P., Jay, F., Arnold, B., Dharmasiri, N., Estelle, M., Voinnet, O. and Jones, J.D. (2006) A plant miRNA contributes to antibacterial resistance by repressing auxin signaling. *Science* 312, 436–439.
- Navarro, L., Zipfel, C., Rowland, O., Keller, I., Robatzek, S., Boller, T. and Jones, J.D.G. (2004) The transcriptional innate immune response to flg22. Interplay and overlap with *Avr* gene-dependent defense responses and bacterial pathogenesis. *Plant Physiol.* 135, 1113–1128.
- Nelson, A.J. and Bushnell, W.R. (1997) Transient expression of anthocyanin genes in barley epidermal cells – potential for use in evaluation of disease response genes. *Transgenic Res.* 6, 233–244.
- Nielsen, K., Olsen, O. and Oliver, R. (1999) A transient expression system to assay putative antifungal genes on powdery mildew infected barley leaves. *Physiol. Mol. Plant Pathol.* 54, 1–12.
- Noel, L., Moores, T.L., van Der Biezen, E.A., Parniske, M., Daniels, M.J., Parker, J.E. and Jones, J.D. (1999) Pronounced intraspecific haplotype divergence at the *RPP5* complex disease resistance locus of *Arabidopsis*. *Plant Cell* 11, 2099–2112.
- Nsarellah, N., Amri, A., Nachit, M.M., El Bouhssini, M. and Lhaloui, S. (2003) New durum wheat with Hessian fly resistance from *Triticum araraticum* and *T. carthlicum* in Morocco. *Plant Breed.* 122, 435–437.
- Oliver, R.E., Cai, X., Xu, S.S., Chen, X. and Stack, R.W. (2005) Wheat-alien species derivatives: a novel source of resistance to fusarium head blight in wheat. *Crop Sci.* 45, 1353–1360.
- Panstruga, R. (2004) A golden shot: how ballistic single cell transformation boosts the molecular analysis of cereal-mildew interactions. *Mol. Plant Pathol.* 5, 141–148.
- Parry, D.W., Jenkinson, P. and McLeod, L. (1995) Fusarium ear blight (scab) in small grain cereals – a review. *Plant Pathol.* 44, 207–238.
- Pós, V., Halász, K., Mesterházy, Á., Csősz, L., Manninger, K., Hunyadi-Gulyás, É., Medzihradský, K., Juhász, T. and Lukács, N. (2005) Proteomic investigation of wheat intercellular washing fluid. *Acta Biol. Szeged* 49, 31–32.
- Puthoff, D.P., Sardesai, N., Subramanyam, S., Nemacheck, J.A. and Williams, C.E. (2005) *Hfr-2*, a wheat cytolytic toxin-like gene, is up-regulated by virulent Hessian fly larval feeding. *Mol. Plant Pathol.* 6, 411–423.
- Rampitsch, C., Bykova, N.V., McCallum, B., Beimcik, E. and Ens, W. (2006) Analysis of the wheat and *Puccinia triticina* (leaf rust) proteomes during a susceptible host-pathogen interaction. *Proteomics* 6, 1897–1907.
- Rockman, M.V. and Kruglyak, L. (2006) Genetics of global gene expression. *Nat. Rev. Genet.* 7, 862–872.
- Roelfs, A.P. (1985) Wheat and rye stem rust. In: A.P. Roelfs and W.R. Bushnell (Eds.), *The Cereal Rusts*, Vol. 2. Academic Press, Orlando, FL, pp. 3–37.
- Rubiales, D., Carver, T.W.L. and Martin, A. (2001) Expression of resistance to *Blumeria graminis* f. sp. *tritici* in 'Chinese Spring' wheat addition lines containing chromosomes from *Hordeum vulgare* and *H. chilense*. *Hereditas* 134, 53–57.
- Ruge, B., Linz, A., Pickering, R., Proeseler, G., Greif, P. and Wehling, P. (2003) Mapping of *Rym14*(Hb), a gene introgressed from *Hordeum bulbosum* and conferring resistance to *BaMMV* and *BaYMV* in barley. *Theor. Appl. Genet.* 107, 965–971.
- Sardesai, N., Nemacheck, J., Subramanyam, S. and Williams, C. (2005) Identification and mapping of *H32*, a new wheat gene conferring resistance to Hessian fly. *Theor. Appl. Genet.* 111, 1167–1173.
- Schultheiss, H., Dechert, C., Kogel, K.-H. and Huckelhoven, R. (2002) A small GTP-binding host protein is required for entry of powdery mildew fungus into epidermal cells of barley. *Plant Physiol.* 128, 1447–1454.

- Schultheiss, H., Dechert, C., Kogel, K.-H. and Huckelhoven, R. (2003) Functional analysis of barley RAC/ROP G-protein family members in susceptibility to the powdery mildew fungus. *Plant J.* 36, 589–601.
- Schultheiss, H., Hensel, G., Imani, J., Broeders, S., Sonnewald, U., Kogel, K.H., Kumlehn, J. and Huckelhoven, R. (2005) Ectopic expression of constitutively activated RACB in barley enhances susceptibility to powdery mildew and abiotic stress. *Plant Physiol.* 139, 353–362.
- Schweizer, P. (2007) Nonhost resistance of plants to powdery mildew—New opportunities to unravel the mystery. *Physiol. Mol. Plant Pathol.* 70, 3–7.
- Schweizer, P., Christoffel, A. and Dudler, R. (1999a) Transient expression of members of the germin-like gene family in epidermal cells of wheat confers disease resistance. *Plant J.* 20, 541–552.
- Schweizer, P., Pokorny, J., Abderhalden, O. and Dudler, R. (1999b) A transient assay system for the functional assessment of defense-related genes in wheat. *Mol. Plant Microbe Interact.* 12, 647–654.
- Schweizer, P., Pokorny, J., Schulze-Lefert, P. and Dudler, R. (2000) Doublestranded RNA interferes with gene function at the single-cell level in cereals. *Plant J.* 24, 895–903.
- Scofield, S.R., Huang, L., Brandt, A.S. and Gill, B.S. (2005) Development of a virus-induced gene-silencing system for hexaploid wheat and its use in functional analysis of the *Lr21*-mediated leaf rust resistance pathway. *Plant Physiol.* 138, 2165–2173.
- Seiffert, U. and Schweizer, P. (2005) A pattern recognition tool for quantitative analysis of in planta hyphal growth of powdery mildew fungi. *Mol. Plant Microbe Interact.* 18, 906–912.
- Serenius, M., Mironenko, N. and Manninen, O. (2005) Genetic variation, occurrence of mating types and different forms of *Pyrenophora teres* causing net blotch of barley in Finland. *Mycol. Res.* 109, 809–817.
- Sharma, H., Ohm, H., Goulart, L., Lister, R., Appels, R. and Benlhabib, O. (1995) Introgression and characterization of barley yellow dwarf virus resistance from *Thinopyrum intermedium* into wheat. *Genome* 38, 406–413.
- Shen, Q., Bieri, S., Zhou, F., Haizel, T., Shirasu, K. and Schulze-Lefert, P. (2003) Recognition specificity and RAR1/SGT1 dependency in barley *Mla* disease resistance alleles to the powdery mildew fungus. *Plant Cell* 15, 732–744.
- Shen, Q.-H., Saijo, Y., Mauch, S., Biskup, C., Bieri, S., Keller, B., Seki, H., Ulker, B., Somssich, I.E. and Schulze-Lefert, P. (2007) Nuclear activity of MLA immune receptors links isolate-specific and basal disease-resistance responses. *Science* 315, 1098–1103.
- Shirasu, K., Lahaye, T., Tan, M.W., Zhou, F., Azevedo, C. and Schulze-Lefert, P. (1999a) A novel class of eukaryotic zinc-binding proteins is required for disease resistance signaling in barley and development in *C. elegans*. *Cell* 99, 355–366.
- Shirasu, K., Nielsen, K., Piffanelli, P., Oliver, R.P. and Schulze-Lefert, P. (1999b) Cell-autonomous complementation of *mlo* resistance using a biolistic transient expression system. *Plant J.* 17, 293–299.
- Shirasu, K. and Schulze-Lefert, P. (2003) Complex formation, promiscuity and multi-functionality: protein interactions in disease-resistance pathways. *Trends Plant Sci.* 8, 252–258.
- Sip, V., Sirlova, L. and Chrpova, J. (2006) Screening for barley yellow dwarf virus-resistant barley genotypes by assessment of virus content in inoculated seedlings. *J. Phytopathol.* 154, 336–342.
- Skadsen, R. and Jing, P. (2008) Transcriptome profile of barley aleurone differs between total and polysomal RNAs: implications for proteome modeling. *Mol. Breed.* 21, 261–269.
- Solomon, P.S., Lowe, R.G.T., Tan, K.-C., Walters, O.D.C. and Oliver, R.P. (2006) *Stagonospora nodorum*: cause of stagonospora nodorum blotch of wheat. *Mol. Plant Pathol.* 7, 147–156.
- Sreenivasulu, N., Altschmied, L., Panitz, R., Hähnel, U., Michalek, W., Weschke, W. and Wobus, U. (2001) Identification of genes specifically expressed in maternal and filial tissues of barley caryopses: a cDNA array analysis. *Mol. Genet. Genomics* 266, 758–767.

- Srichumpa, P., Brunner, S., Keller, B. and Yahiaoui, N. (2005) Allelic series of four powdery mildew resistance genes at the *Pm3* locus in hexaploid bread wheat. *Plant Physiol.* 139, 885–895.
- Staal, J., Kaliff, M., Bohman, S. and Dixelius, C. (2006) Transgressive segregation reveals two *Arabidopsis* TIR-NB-LRR resistance genes effective against *Leptosphaeria maculans*, causal agent of blackleg disease. *Plant J.* 46, 218–230.
- Stokstad, E. (2007) Plant pathology: deadly wheat fungus threatens world's breadbaskets. *Science* 315, 1786–1787.
- Subramanyam, S., Sardesai, N., Puthoff, D.P., Meyer, J.M., Nemacheck, J.A., Gonzalo, M. and Williams, C.E. (2006) Expression of two wheat defense-response genes, Hfr-1 and Wci-1, under biotic and abiotic stresses. *Plant Sci.* 170, 90–103.
- Tai, Y.-S. and Bragg, J. (2007) Dual applications of a virus vector for studies of wheat-fungal interactions. *Biotechnology* 6, 288–291.
- Tang, S.X., Zhuang, J.J., Wen, Y.X., Ai, S.J.A., Li, H.J. and Xu, J. (1997) Identification of introgressed segments conferring disease resistance in a tetrageneric hybrid of *Triticum*, *Secale*, *Thinopyrum*, and *Avena*. *Genome* 40, 99–103.
- Tavernarakis, N., Wang, S.L., Dorovkov, M., Ryazanov, A. and Driscoll, M. (2000) Heritable and inducible genetic interference by double-stranded RNA encoded by transgenes. *Nat. Genet.* 24, 180–183.
- Thordal-Christensen, H. (2003) Fresh insights into processes of nonhost resistance. *Curr. Opin. Plant Biol.* 6, 351–357.
- Tooker, J.F. and De Moraes, C.M. (2007) Feeding by Hessian fly [*Mayetiola destructor* (Say)] larvae does not induce plant indirect defences. *Ecol. Entomol.* 32, 153–161.
- Tosa, Y. (1989) Genetic analysis of the avirulence of wheatgrass powdery mildew fungus on common wheat. *Genome* 32, 913–917.
- Trail, F., Xu, J.-R., Miguel, P.S., Halgren, R.G. and Corby Kistler, H. (2003) Analysis of expressed sequence tags from *Gibberella zeae* (anamorph *Fusarium graminearum*). *Fungal Genet. Biol.* 38, 187–197.
- Trujillo, M., Altschmied, L., Schweizer, P., Kogel, K.H. and Huckelhoven, R. (2006) Respiratory Burst Oxidase Homologue A of barley contributes to penetration by the powdery mildew fungus *Blumeria graminis* f. sp. *hordei*. *J. Exp. Bot.* 57, 3781–3791.
- Trujillo, M., Troeger, M., Niks, R.E., Kogel, K.H. and Huckelhoven, R. (2004) Mechanistic and genetic overlap of barley host and non-host resistance to *Blumeria graminis*. *Mol. Plant Pathol.* 5, 389–396.
- Wang, Y., Yang, L., Xu, H., Li, Q., Ma, Z. and Chu, C. (2005) Differential proteomic analysis of proteins in wheat spikes induced by *Fusarium graminearum*. *Proteomics* 5, 4496–4503.
- Wanyera, R., Kinyua, M.G., Jin, Y. and Singh, R.P. (2006) The spread of stem rust caused by *Puccinia graminis* f. sp. *tritici*, with virulence on *Sr31* in wheat in eastern Africa. *Plant Dis.* 90, 113.
- Wei, F., Gobelman-Werner, K., Morroll, S.M., Kurth, J., Mao, L., Wing, R., Leister, D., Schulze-Lefert, P. and Wise, R.P. (1999) The *Mla* (powdery mildew) resistance cluster is associated with three NBS-LRR gene families and suppressed recombination within a 240-kb DNA interval on chromosome 5S (1HS) of barley. *Genetics* 153, 1929–1948.
- Wei, F., Wing, R.A. and Wise, R.P. (2002) Genome dynamics and evolution of the *Mla* (powdery mildew) resistance locus in barley. *Plant Cell* 14, 1903–1917.
- Williams, C.E., Collier, C.C., Nemacheck, J.A., Liang, C. and Cambron, S.E. (2002) A lectin-like wheat gene responds systemically to attempted feeding by avirulent first-instar Hessian fly larvae. *J. Chem. Ecol.* 28, 1411–1428.
- Williams, K., Willshire, K., Olson, S., Matic, M. and Kuchel, H. (2006) Mapping of a novel QTL for resistance to cereal cyst nematode in wheat. *Theor. Appl. Genet.* 112, 1480–1486.
- Williams, R.B.H., Chan, E.K.F., Cowley, M.J. and Little, P.F.R. (2007) The influence of genetic variation on gene expression. *Genome Res.* 17, 1707–1716.

- Wise, R.P., Caldo, R.A., Hong, L., Shen, L., Cannon, E.K. and Dickerson, J.A. (2007a) BarleyBase/PLEXdb: a unified expression profiling database for plants and plant pathogens. In: D. Edwards (Ed.), *Methods in Molecular Biology*, Vol. 406, *Plant Bioinformatics – Methods and Protocols*. Humana Press, Totowa, NJ, pp. 347–363.
- Wise, R.P., Moscou, M.J., Bogdanove, A.J. and Whitham, S.A. (2007b) Transcript profiling in host-pathogen interactions. *Annu. Rev. Phytopathol.* 45, 329–369.
- Yahiaoui, N., Brunner, S. and Keller, B. (2006) Rapid generation of new powdery mildew resistance genes after wheat domestication. *Plant J.* 47, 85–98.
- Yahiaoui, N., Srichumpa, P., Dudler, R. and Keller, B. (2004) Genome analysis at different ploidy levels allows cloning of the powdery mildew resistance gene *Pm3b* from hexaploid wheat. *Plant J.* 37, 528–538.
- Yao, Y., Guo, G., Ni, Z., Sunkar, R., Du, J., Zhu, J.-K. and Sun, Q. (2007) Cloning and characterization of microRNAs from wheat (*Triticum aestivum* L.). *Genome Biol.* 8, R96.
- Zakhrabekova, S., Gamini Kannangara, C., von Wettstein, D. and Hansson, M. (2002) A microarray approach for identifying mutated genes. *Plant Physiol. Biochem.* 40, 189–197.
- Zambino, P.J., Kubelik, A.R. and Szabo, L.J. (2000) Gene action and linkage of avirulence genes to DNA markers in the rust fungus *Puccinia graminis*. *Phytopathology* 90, 819–826.
- Zhang, L., Fetch, T., Nirmala, J., Schmierer, D., Brueggeman, R., Steffenson, B. and Kleinhofs, A. (2006) *Rpr1*, a gene required for *Rpg1*-dependent resistance to stem rust in barley. *Theor. Appl. Genet.* 113, 847–855.
- Zhao, H., Liu, X. and Chen, M.S. (2006) *H22*, a major resistance gene to the Hessian fly (*Mayetiola destructor*), is mapped to the distal region of wheat chromosome 1DS. *Theor. Appl. Genet.* 113, 1491–1496.
- Zhou, F., Kurth, J., Wei, F., Elliott, C., Vale, G., Yahiaoui, N., Keller, B., Somerville, S., Wise, R. and Schulze-Lefert, P. (2001) Cell-autonomous expression of barley *Mla1* confers race-specific resistance to the powdery mildew fungus via a *Rar1*-independent signaling pathway. *Plant Cell* 13, 337–350.
- Zhou, H., Li, S., Deng, Z., Wang, X., Chen, T., Zhang, J., Chen, S., Ling, H., Zhang, A., Wang, D. and Zhang, X. (2007) Molecular analysis of three new receptor-like kinase genes from hexaploid wheat and evidence for their participation in the wheat hypersensitive response to stripe rust fungus infection. *Plant J.* 52, 420–434.
- Zhou, W., Kolb, F.L. and Riechers, D.E. (2005) Identification of proteins induced or upregulated by Fusarium head blight infection in the spikes of hexaploid wheat (*Triticum aestivum*). *Genome* 48, 770–780.
- Zimmermann, G., Baumlein, H., Mock, H.P., Himmelbach, A. and Schweizer, P. (2006) The multigene family encoding germin-like proteins of barley. Regulation and function in basal host resistance. *Plant Physiol.* 142, 181–192.

Chapter 20

Developmental and Reproductive Traits in the Triticeae

David A. Laurie

Abstract This chapter reviews what is known about genes that control the timing of flowering during the year or inflorescence development in the Triticeae, focusing on barley and wheat. The methods used to identify major genes controlling flowering in response to extended periods of low temperature (*Vernalization*; *Vrn* genes) or day length (*Photoperiod*; *Ppd* genes) and genes controlling inflorescence development are considered. This shows that direct candidate gene approaches have proved less effective than positional cloning or composite fine mapping/candidate gene methods. The implications for identifying new target genes are discussed. The analysis of genetic pathways identified in models is also considered as an aid to understanding how genes in the Triticeae affect phenotypic variation. Finally, the evolution of these traits under domestication is considered in relation to the finding that orthologous genes and similar mutations are found to be major contributors to adaptive variation in different species.

20.1 Introduction

This chapter will cover two aspects of reproduction. The first is the genetic control of the timing of flowering during the year. The second is the genetic control of inflorescence development.

The timing of flowering during the year is an important adaptive trait in plants because it plays a key role in reproductive success. Flowering needs to occur when conditions for pollination, seed development and seed dispersal are optimal. For most plants this means that flowering is restricted to a specific time of year. For crops the situation is particularly interesting from an evolutionary perspective because domestication has resulted in the diversification of species

D.A. Laurie (✉)

Department of Crop Genetics, John Innes Centre, Norwich Research Park,
NR4 7UH, UK

e-mail: david.laurie@bbsrc.ac.uk

into 'agritypes' that are adapted to a wide range of environments and farming practices. The distinction between winter (fall/autumn sown) and spring (spring sown) cereals is one obvious case. Selection has clearly acted to allow the best use of the available growing season and hence to maximize productivity in individual environments.

From this, an obvious question to ask is how does a plant determine when to flower? The mechanisms vary between species and include factors such as plant size, nutrient status and water availability. Importantly, many plants also exploit reliable cues from the environment, particularly temperature and day length (photoperiod). The wild ancestors of the Triticeae crops, and many domestic varieties, have a winter annual life style. That is, they germinate in autumn and overwinter in a vegetative state before committing to flowering in response to the long days experienced in spring. The genetic systems determining the need for an extended period of cold before flowering (vernalization) and the response to photoperiod have both been altered by selection during domestication. The genetic basis of these changes is now becoming understood (recently reviewed by Cockram et al. 2007a).

The timing of flowering is, however, only one aspect of reproduction. The plant must also produce seeds in sufficient quantity for survival in the wild and for acceptable yield in agriculture. The development of the seed bearing structure, the inflorescence, is therefore critical.

Studying the genetic basis of development is important because it allows us to understand fundamental aspects of crop growth and the differences that are needed to allow productivity in different agricultural situations. To understand this fully we need to know the identities of the underlying genes, how they act and interact, and what range of alleles are available. Once the genes are known the directed manipulation of the plant becomes possible. This can be by marker-assisted breeding, by the identification and introgression of novel alleles from germplasm collections, by the creation of new alleles by mutation or by genetic engineering.

The major crop species in the Triticeae (wheat, barley and rye) have individual advantages and disadvantages for analyzing reproductive traits. Barley (*Hordeum vulgare*), rye (*Secale cereale*) and einkorn wheat (*T. monococcum*) are diploid ($2n = 2x = 14$) and are suited to the detection of recessive mutations. Tetraploid pasta (*T. durum*) and hexaploid bread (*T. aestivum*) wheats are tolerant of aneuploidy and have been used to create a range of cytogenetic stocks. This is assisted by the fact that tetraploid and hexaploid types behave genetically as diploids (i.e. there is normally no recombination between the homoeologous chromosomes; see Chapter 8) but because of the close relationship of the genomes there is a strong tendency for recessive mutations on one genome to be concealed by functional genes on the others. The study of *waxy* mutations is a good example (Slade et al. 2005).

Numerous cytogenetic stocks in which individual chromosomes or chromosome segments are missing or replaced exist in wheat. These have proved extremely valuable for assigning genetic markers to particular chromosomes and are now being used for flow sorting of chromosomes for physical mapping

(see Chapter 10). Historically, aneuploid lines were used to assign genes to chromosomes and this allowed major *Photoperiod* (*Ppd*) and *Vernalization* (*Vrn*) genes to be located to individual group 2 and group 5 chromosomes, respectively, prior to the development of marker technologies (reviewed by Law and Worland 1997). Subsequently, wheat genetics developed slowly. There are several reasons for this. One was the difficulty of developing comprehensive genetic maps because of a lack of markers and a low level of polymorphism compared to species such as maize (*Zea mays*). A second was the large genome size. This made the construction of genomic libraries technically challenging and fostered the belief that cloning genes directly from wheat, barley or rye was too difficult. Technically facile model species, particularly *Arabidopsis*, attracted many researchers because of their extensive genetic and genomic resources (including the first complete plant genome sequence) and a widely held belief that knowledge from models would be directly applicable to crop species.

Recently, several factors have altered this situation. Firstly, funding became available for the construction of high quality Triticeae resources including BAC libraries and extensive EST collections (Chapter 9). Secondly, colinearity with rice (*Oryza sativa*), established in the early 1990s (Moore et al. 1993, 1995), became much more valuable as the complete rice genome sequence emerged. More recently still, *Brachypodium distachyon* has become established as a temperate grass model and a draft release of its genomic sequence was made in 2008 (<http://www.brachypodium.org/>). This facilitates the development of large numbers of genetic markers targeted at specific regions and allows the gene content of individual segments of crop genomes to be predicted, permitting candidate genes to be identified more readily.

These advances now enable researchers to identify genes directly in Triticeae species themselves and to take much greater advantage of the extensive analyses of reproductive traits that have been carried out in model species such as *Arabidopsis* and, more recently, rice (see Bäurle and Dean 2006; Schmitz and Amasino 2007; Kobayashi and Weigel 2007 for reviews of flowering and Bommert et al. 2005; Itoh et al. 2006; Benlloch et al. 2007 for reviews of inflorescence development).

Inflorescence architecture in the grasses shows striking variation resulting from differences in branching and the organization of spikelets and florets (flowers). Figure 20.1 shows an outline of the differences between rice, maize, barley and wheat. Barley and wheat inflorescences do not have lateral branches and form spikelets on a central rachis. Although closely related, they show interesting differences in inflorescence organization. In wild type (2-rowed) barley the spikelets are determinate with each node bearing a single central floret flanked by two partially developed lateral florets. Cultivated barley also has a six-rowed form in which all three florets develop fully. The barley inflorescence is indeterminate: that is, it does not terminate with the formation of a flower or other defined structure. In wheat, the spikelet is indeterminate and several florets are borne at each rachis node. The inflorescence is determinate and ends in a terminal spikelet. The genetic basis for

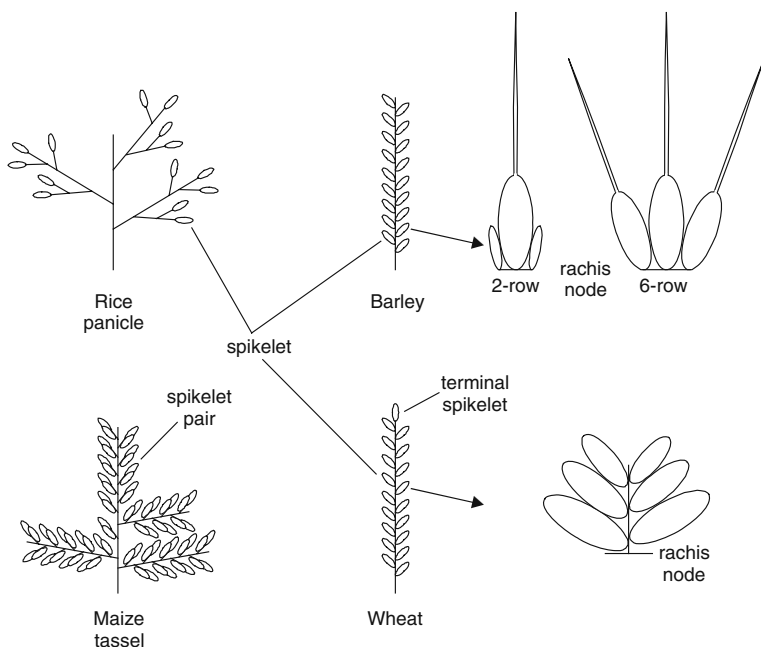


Fig. 20.1 Variation in inflorescence architecture between rice, maize, barley and wheat arising from differences in branching patterns, spikelet organization and determinacy

this difference between barley and wheat is unknown. Most information on inflorescence development currently comes from maize and rice where numerous mutations affecting branching and meristem fate have been identified (see Bommert et al. 2005; Malcomber et al. 2006; Kellogg 2007 for reviews). Interestingly, this reveals differences in genetic mechanism even within the grasses. For example, the *ramosa1*, 2 and 3 mutations of maize affect lateral branching. All three genes have been cloned but only *ramosa2* is conserved outside the Andropogoneae (maize, millet, sorghum group) (Kellogg 2007). In the case of *ramosa3* (a trehalose-6-phosphate phosphatase enzyme; Satoh-Nagasawa et al. 2006) a closely related duplicate gene is present in other species.

20.2 Gene Catalogues

Flowering time and development mutations, both natural and induced, are well known in the Triticeae. Specific examples are described below and comprehensive lists are given in the Barley Genetics Newsletter Vol. 26 (1996) and the Wheat Gene Catalogue at the GrainGenes web site (<http://wheat.pw.usda.gov/GG2/index.shtml>).

20.3 Identifying Flowering Time Genes in the Triticeae

Several major genes in wheat and barley have been the targets of cloning efforts, and in considering how best to approach future work it is useful to consider the methods used. These can be divided into the broad categories of candidate gene, positional cloning and positional/candidate hybrids.

20.3.1 *The Candidate Gene Method*

Although not a flowering time gene, the cloning of the *Reduced Height-1 (Rht-1)* gene is an instructive example of this approach. The gene was first identified in *Arabidopsis* as a member of the GRAS family of proteins with a distinctive N-terminal region containing a conserved five amino-acid DELLA motif (Peng et al. 1997). The protein acts as a growth repressor whose action is relieved by the gibberellic-acid (GA) mediated destruction of the protein. The wheat gene was cloned by PCR using the *Arabidopsis* gene sequence and corresponding rice ESTs. Loss of the N-terminal region removes the GA regulation and produces a dominant GA-insensitive dwarf phenotype in wheat, maize and barley while loss of function mutations in barley and rice give the *slender* mutant phenotype (Peng et al. 1999; Ikeda et al. 2001; Chandler et al. 2002). Recently DELLA proteins have been implicated in the control of growth via other signaling systems, including auxins and abiotic stresses, suggesting that they act as integrators of growth and stress related pathways and are likely to impact on flowering (Fu and Harberd 2003; Achard et al. 2006).

This example is important for two reasons. Firstly, the idea that *GAI* and *Rht* might be related was not based on dwarf phenotype alone but on physiological data on gibberellic acid (GA) levels and a similar response to the GA biosynthesis inhibitor paclobutrazol, stressing the importance of an in depth understanding of traits before the selection of candidate genes. Secondly, this example dates from 1999 but the direct use of candidate genes has not subsequently established itself as the method of choice for identifying targets in the Triticeae. One reason for this is that although many genetic pathways are evolutionarily conserved, they are not so conserved as to allow the prediction of the individual pathway component that provides trait variation, especially in the absence of detailed comparative physiology. In addition, even conserved pathways have novel components and a significant problem is caused by differences in the evolution of gene families. This results from the duplication and deletion of individual genes and from differences in the whole genome duplications that have occurred in the monocot and dicot lineages (Yu et al. 2005; Blanc et al. 2003).

20.3.2 The Positional Cloning Method

The isolation of the wheat *Vernalization-2* (*Vrn-2*) gene by Yan et al. (2004a) used a classical positional cloning approach which involved the identification of tightly-linked flanking markers and the sequencing of the intervening region. Genomic model species played a key role as colinearity with rice provided closely linked markers. *Vrn-2* was shown to be a member of the *CO*, *COL1* and *TOC1* (CCT) domain family, distantly related to the Arabidopsis *CONSTANS* (*CO*) gene that is involved in photoperiod response (Fig. 20.2).

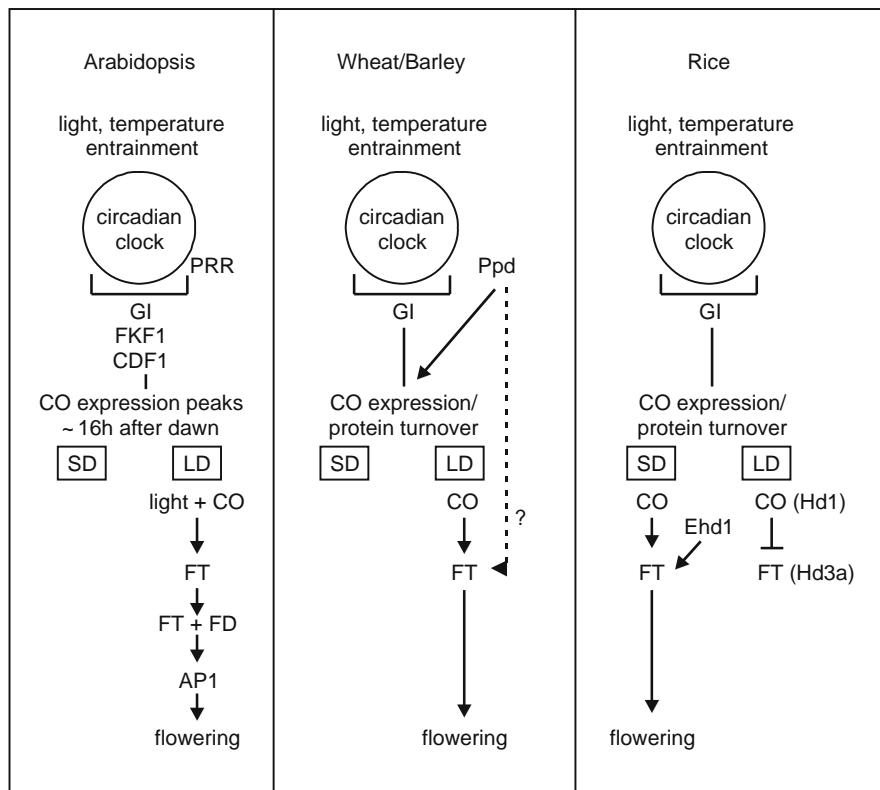


Fig. 20.2 Summary of the pathway regulating photoperiod response in Arabidopsis, Triticeae species and rice. Drawn from Kobayashi and Weigel (2007); Beales et al. (2007) and Doi et al. (2004). Key features are the circadian regulation of *CONSTANS* (*CO*) and the coincidence of its expression peak with light which occurs in long days (LD) but not short days (SD). *CO* protein is active in the light period, promoting the expression of *FT*. There is good evidence that *FT* protein is a mobile signal (the mysterious ‘florigen’ of classical plant physiology) linking day length perception in the leaf with the shoot apex where *FT* and *FD* interact to regulate meristem differentiation (reviewed by Kobayashi and Weigel 2007). While the core photoperiod pathway is well conserved, additional genes have been recruited in different evolutionary lineages. For example, the *Ehd1* gene, which can activate *FT* in an *Hd1* (rice *CO*) independent manner, is only known in rice (Doi et al. 2004). Gene names are given in the text

In the diploid (*T. monococcum*) *Vrn-2* is a member of a tandem pair of CCT domain genes while in pasta wheat (*T. durum*) and probably barley there are three. The spring type allele in *T. monococcum* results from a mutation in the CCT domain of one gene that alters a highly conserved amino-acid also affected in the Arabidopsis *co7* mutation. It is not known what role in flowering, if any, the associated CCT domain gene(s) may have. The equivalent region of barley has been sequenced from a spring cultivar which shows a complete deletion of the ZCCT gene cluster (Yan et al. 2004a). *Vrn-2* is a strong repressor of flowering and loss of function is sufficient to remove the requirement for vernalization. Variation at *Vrn-2* is not a known cause of winter/spring differences in tetraploid or hexaploid wheat, probably because homozygosity for loss of function would be required on each genome for the phenotype to be seen.

20.3.3 The Positional Cloning/Candidate Gene Hybrid Method

This has been the approach for three other flowering time genes that have been cloned. In the case of the *Vernalization-1* (*Vrn-1*) gene (Yan et al. 2003), fine mapping identified a region containing a MADS-box family transcription factor related to the *APETALA1* (*API*) and *FRUITFULL* (*FUL*) genes of Arabidopsis that have roles in determining meristem identity and the transition from vegetative to reproductive growth. The gene was also identified by Danyluk et al. (2003) and Murai et al. (2003). Confirmation that the candidate was *Vrn-1* was obtained by analysis of gene expression, by sequencing alleles from known spring and winter genotypes (Yan et al. 2003) and by transformation tests (Loukoianov et al. 2005). Mutations from the wild type winter habit to spring habit are due to altered regulation arising from insertion or deletion events in the promoter or first intron (wheat – Yan et al. 2004b; Beales et al. 2005; Fu et al. 2005; barley – von Zitzewitz et al. 2005; Cockram et al. 2007b). Preston and Kellogg (2007) show that *Vrn-1* and its paralogue *Ful-2* are likely to have multiple redundant functions in early inflorescence development. However, complete loss of function mutations of *Vrn-1* identified in *T. monococcum* remain vegetative, showing that this gene has an essential role in the transition to flowering (Shitsukawa et al. 2007).

In contrast to *Vrn-2*, mutations to spring habit at *Vrn-1* are incompletely dominant and are effective in diploid or polyploid settings. Spring wheats from different regions differ in which gene is mutated (Iwaki et al. 2000), but it is not clear if this reflects historical patterns in the origin of independent mutations or if different mutations have selective advantages in specific environments. Analysis of spring barley from Europe shows at least six mutations in intron 1, showing that the evolution of the spring type has occurred repeatedly (Cockram et al. 2007b).

In the case of the *Vernalization-3* (*Vrn-3*) gene of barley and wheat (Yan et al. 2006), fine mapping identified a region containing a cereal homologue of

FLOWERING LOCUS T (FT), a key inducer of flowering in Arabidopsis and the gene underlying the *Heading date 3a* flowering time QTL in rice (Kojima et al. 2002) (Fig. 20.2). The mutant form in wheat had a transposable element inserted in its promoter but the basis of the mutation in barley is not known. Importantly, the cloning of *Vrn-1* and *Vrn-3* differed from a traditional positional cloning approach because no complete physical map of either region was required.

A third gene identified by the hybrid positional/candidate approach was *Photoperiod-1 (Ppd-1)* which controls flowering in response to day length. As for *Vrn-1* and *Vrn-3*, the gene was first identified in a diploid, in this case barley, where a recessive mutation causes late flowering in long days but does not affect flowering in short days. Fine scale mapping exploiting colinearity with rice and *Brachypodium* identified a region containing a member of the *Pseudo-Response Regulator (PRR)* family which is another group of CCT domain genes. In Arabidopsis, *PRR* genes are part of, or are closely associated with, the circadian clock (Mizuno and Nakamichi 2005), making this an excellent candidate for photoperiod response (Fig. 20.2). Sequencing of mutant and wild type alleles showed that the barley mutation is likely to produce a non-functional protein because of an alteration to a highly conserved amino-acid in the CCT domain (Turner et al. 2005).

In wheat and barley, the selection of spring growth habit has resulted from similar mutations in the *Vrn-1* and *Vrn-2* genes. In the case of the *Ppd* genes the situation is more complex because barley and wheat have contrasting mutant phenotypes. In barley, a recessive mutation confers late flowering in long days. In wheat semi-dominant mutations confer early flowering in short or long days which is referred to as a photoperiod insensitive phenotype. Therefore, although the barley and wheat mutations mapped in colinear positions on the short arms of the group 2 chromosomes (2H, 2B and 2D), there was no definite expectation that the same gene was involved. However, sequencing the *PRR* genes from wheat varieties with known *Ppd* alleles showed that this was the gene (Beales et al. 2007). In the *Ppd-D1a* mutation, early flowering was associated with deletion of a 2 kb region upstream of the gene and resulting misexpression of the transcript. Specifically, the normally tight circadian control of *PRR* expression (rising to a peak about 6 h after dawn and returning to very low levels by 12 h) was altered and the gene was expressed throughout the 24 h period. Thus, contrasting early and late flowering phenotypes in wheat and barley can be selected by using alternative gain or loss of function mutations in orthologous genes.

More recently, analysis of photoperiod insensitive mutations in durum wheat (*T. durum*) identified two deletions in the A genome *PRR* gene that remove a region common to the *Ppd-D1a* allele (Wilhelm, Turner and Laurie, unpublished). A remaining puzzle in wheat is the nature of the photoperiod insensitive mutation on the B genome. Early flowering alleles are well known on the B genome and although the 2B *PRR* gene cosegregates with flowering phenotype no candidate mutation has been identified so far (Beales et al.

2007). This may mean that another gene is involved or, more likely, that a different mutational mechanism is involved that does not involve deletion of the region affected in the A and D genome alleles.

20.4 Identifying Inflorescence Development Genes in the Triticeae

Numerous mutations affecting inflorescence development are known in the Triticeae but, to date, only three have been cloned. As with flowering time, several approaches have been used.

20.4.1 The Candidate Gene Method

The *Hooded* mutation, which affects the development of the lemma, was identified by a candidate gene approach as a homologue of the maize homeodomain gene *Knotted1* (Müller et al. 1995). Dominant *Knotted1* mutations cause alterations in the base-to-tip patterning of leaves (Vollbrecht et al. 1991) while dominant mutations in the barley *Hooded* gene specifically alter the developing lemma, causing a second flower to develop in an inverted orientation. *Hooded* results from a sequence duplication in intron 4 that causes the gene to be ectopically expressed in the lemma, leading to the formation of a new flower. To date, no other Triticeae mutations have been identified directly using genes from maize. Improved map locations of barley mutations, together with more refined colinearity with genome sequences, may in future lead to the identification of further genes (Rossini et al. 2006). This becomes a positional cloning/candidate gene hybrid approach.

20.4.2 The Positional Cloning Method

At each node in the inflorescence, wild type (two-rowed) barley has a single fully formed flower (floret) flanked by partially developed lateral florets. The recessive *vrs1* mutation causes fully developed flowers to form in the lateral positions, giving a six-rowed inflorescence (Fig. 20.1). The *Vrs1* gene was identified by a positional cloning approach, again exploiting colinearity with rice to develop targeted closely linked markers, although in this case colinearity was disrupted by transposition of a chromosome segment (Pourkheirandish et al. 2007). *Vrs1* is a member of the homeodomain-leucine zipper I-class homeobox family (Komatsuda et al. 2007). Sequencing from different varieties shows that the six-row phenotype, caused by loss of function of *vrs-1*, has arisen several times independently during the domestication of barley.

The *Q* mutation of wheat makes it easy to separate the seed from the rest of the flower at maturity. This is an important character associated with

domestication. Using a combination of high resolution mapping, induced mutations and chromosome walking using a *T. monococcum* BAC library, Faris et al. (2003) and Simons et al. (2006) identified *Q* as a transcription factor of the *APETALA2* (*AP2*) family. *Q* is highly homologous to the *indeterminate spikelet1* (*ids1*) gene of maize which affects spikelet meristem fate (Chuck et al. 1998). In this case, the difference between the wheat and maize phenotypes meant that *ids1* was not an obvious candidate gene. Simons et al. (2006) showed that the mutant allele of *Q* is more abundantly transcribed than the wild type gene but suggest that altered protein/protein interaction is the basis of the mutant (domestic) phenotype. However, Chuck et al. (2007) identified *tassel-seed4* in maize as a mutation in a microRNA that targets *AP2* genes including *ids1*. They note that its target sequence is mutated in the *Q* gene, suggesting that loss of microRNA mediated regulation is the basis of the wheat phenotype.

20.5 Understanding Gene Function

20.5.1 The Analysis of Genetic Pathways

The above examples show the ways in which developmentally significant genes have been isolated from the Triticeae. This is clearly important, but effective exploitation of these results requires an understanding of the ways in which these genes work. Genetic pathways defined in model species have proved valuable for this by providing conceptual frameworks within which to explore gene function.

Bioinformatic analysis of ESTs and other sequences can rapidly establish whether a pathway is likely to be evolutionarily conserved and identify components for analysis. In the case of photoperiod response, the candidate approach of isolating barley homologues of genes in the Arabidopsis photoperiod pathway failed to identify *Ppd* directly. However, the availability of cereal homologues of other genes in the pathway was extremely valuable for interpreting the way in which *Ppd* mutations affect flowering. Firstly, the existence of cereal homologues suggested the photoperiod pathway was conserved and therefore that *Ppd* would operate by affecting a known developmental pathway that contains orthologues of *GIGANTEA* (*GI*) *CONSTANS* (*CO*) and *FLOWERING LOCUS T* (*FT*) at its core (Griffiths et al. 2003; Dunford et al. 2005; Faure et al. 2007). The induction of *FT* has been shown to be the key trigger in flowering (reviewed by Kobayashi and Weigel 2007) and while the interaction between *CO* and *FT* may vary between plants that flower in short or long days (Fig. 20.2), the induction of *FT* is always associated with the promotion of flowering.

In barley, this process could be studied by analyzing the expression of photoperiod pathway genes in genotypes with wild type and mutant *Ppd* alleles. The *Ppd* mutation did not affect the expression of the *Ppd* gene itself but

was associated with altered expression of the barley *CO* homologue and a reduction in *FT* expression consistent with the late flowering phenotype (Turner et al. 2005).

In wheat, the early flowering photoperiod insensitive mutations show a loss of the normal circadian regulation of the *Ppd* gene so that the gene becomes expressed throughout the day. This is accompanied by strong induction of *FT* in short or long days but the mechanism by which these events are connected is not entirely clear (Beales et al. 2007). The amplitude of *CO* expression is reduced but the current data do not exclude a direct effect on *FT* (dotted line in Fig. 20.2).

In the case of vernalization the situation is more complicated because the mechanism clearly differs from that in *Arabidopsis* where *FRIGIDA* (*FRI*) and *FLC* (a MADS-box transcription factor) are the major determinants of variation. *FRI* upregulates the floral repressor *FLC* early in development and subsequently *FLC* expression is stably reduced by exposure to cold by a mechanism that involves epigenetic changes to its chromatin structure. *FLC* is then reset in the next generation. Mutation to spring (rapid cycling) type in *Arabidopsis* is commonly by loss of *FRI* function (Shindo et al. 2005) and can also occur by variation at *FLC* (Shindo et al. 2006). Comparison of *Arabidopsis* ecotypes shows that the rapid cycling type has arisen repeatedly from independent mutations of *FRI*.

No *FLC* orthologue has been identified in cereals where the major repressor of flowering (*Vrn-2*) is a different class of gene. It also shows different behaviour as *Vrn-2* expression is also reduced in short days (Dubcovsky et al. 2006; Trevaskis et al. 2006). Additional genes such as *Vrt2* (another MADS-box transcription factor) may also be important (Kane et al. 2005, 2007). This means that there is still dispute about how the cereal components fit together into a pathway and three alternatives from different groups are shown in Fig. 20.3. A further important difference is that *Vrn-1* is expressed in the leaf and shoot apex of cereals while *API* in *Arabidopsis* is only expressed in the apex. Preston and Kellogg (2008) suggest distinct roles for the *Vrn-1* gene, one in systemic signaling that induces flowering competence and the other in the control of floral transition in the meristem itself.

Despite these complications, it seems likely that vernalization in cereals parallels the process in *Arabidopsis* in that it regulates flowering by controlling the level of *FT* expression (Hemming et al. 2008). Cereals also possess homologues of several genes involved in *FLC* regulation in *Arabidopsis* including *FCA* and *FY* which are part of the autonomous pathway (Winichayakul et al. 2005). Therefore, there are likely to be parallels in mechanism, including the epigenetic regulation of the vernalization target, even if the target gene is different. In support of this idea, Triticeae species have homologues of *VERNALIZATION INSENSITIVE 3* (*VIN3*) (Fu et al. 2007), an *Arabidopsis* gene that is upregulated at low temperatures and which is implicated in the early stages of cold perception and vernalization (Sung and Amasino 2004).

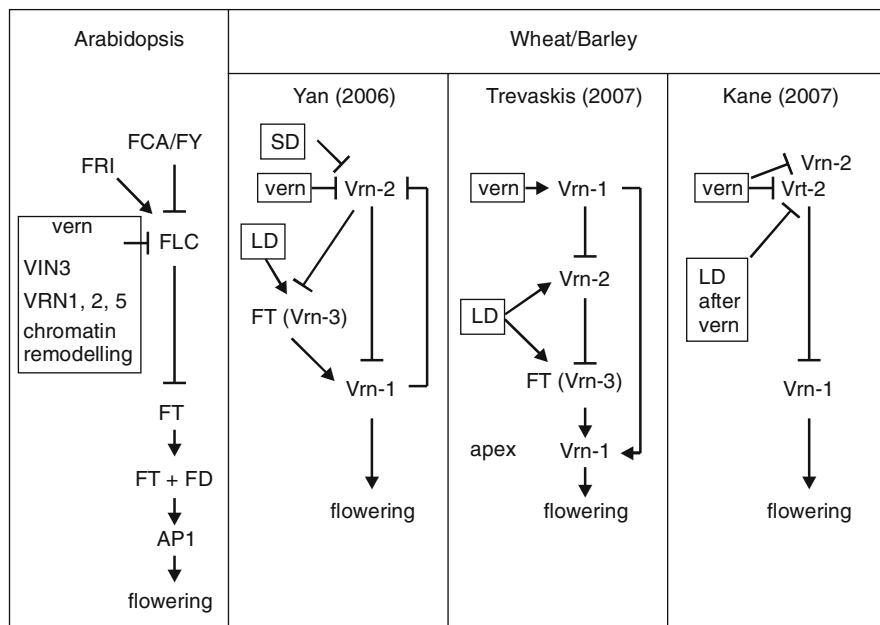


Fig. 20.3 Models of the vernalization pathway and its targets in Arabidopsis (redrawn from Bäurle and Dean (2006) and Schmitz and Amasino (2007)) and temperate cereals. In the latter, models are redrawn from Yan et al. (2006); Trevaskis et al. (2007) and Kane et al. (2007), respectively. Vernalizing temperatures, short days and long days are indicated by vern, SD and LD, respectively. Note: full gene names are given in the text but note that *VRN* in Arabidopsis refers to B3 domain, polycomb-like or PHD finger genes that are unrelated to the cereal *Vrn* genes

20.5.2 Validation of Candidate Flowering Genes

All candidate genes need to be validated and this has become much more straightforward recently because of the development of additional resources in the Triticeae. These include TILLING and the development of more efficient transformation systems (see Chapter 13). Transformation efficiency, however, remains dependent on genotype. In barley the most readily transformable genotype is 'Golden Promise'. In terms of flowering behaviour this is a typical European spring variety with a deletion of the *Vrn-2* gene, a mutant allele of *Vrn-1* and a loss of function mutation of *Ppd-H1*. This would be an advantage for some studies such as the testing of different functional alleles of *Ppd-1*, but is a disadvantage for testing some aspects of vernalization. For example, introducing a putative functional copy of *Vrn-2* would not confer a winter habit unless a wild type *Vrn-1* allele was also introduced. This could be done by crossing wild type alleles into 'Golden Promise' to produce a series of alternative backgrounds with different endogenous flowering characteristics or by improving transformation efficiency in other genotypes.

20.6 Advances in Triticeae Genomics and Gene Identification

The previous experience of isolating genes from barley and wheat enables us to consider the most efficient strategy for new targets. Best practice would seem to be fine scale mapping of the trait using colinearity with existing sequenced genomes (rice and *Brachypodium*) to provide targeted markers. This can rapidly identify a small region for which the sequenced genomes provide candidate genes. This can be verified by small scale physical mapping and sequencing of BAC clones from the Triticeae species of interest. This approach will rapidly become more efficient as physical mapping and sequencing within the Triticeae species themselves develops.

An important observation from the cloning work to date is that several important mutations such as spring alleles of *Vrn-1* in barley and wheat, spring alleles of *Vrn-3*, photoperiod insensitive *Ppd-1* alleles in wheat and *Hooded* in barley are caused by changes in regulatory regions and do not affect protein sequence. Therefore it will be important to develop methods that will allow the cloning and sequencing of genomic regions from varieties other than those for which large insert libraries are developed. In some cases PCR will be sufficient but a more secure method would be the ability to isolate specific large (BAC sized) fragments from any genotype. This is important because examples in maize show that regulatory regions controlling phenotype can lie 40 (Clark et al. 2006) to 70 (Salvi et al. 2007) kb from the coding region of the gene.

The process of gene discovery in the Triticeae will undoubtedly accelerate. Firstly, there are many more markers available for the primary mapping, and an increased use of gene-based markers assists alignment with sequenced genomes. This alignment is much more effective now that additional genomes are being sequenced. The combination of rice and *Brachypodium* is very powerful for Triticeae species. However, no model will ever be as good as the species itself, and the key to fully understanding Triticeae traits is to have genomic sequences of the crops themselves. This is now feasible due to advances in physical mapping and sequencing technologies. Even a physical map would be a huge advance because sequencing could be targeted to regions of interest.

A further important consideration is whether researchers, particularly those starting their careers, will be attracted to the Triticeae by these new opportunities. If this does not occur then it is conceivable that the technical ability to isolate cereal genes will outstrip the number of well defined biological targets and the number of research groups able to identify them.

Once genes are known it is possible to develop allele specific markers that allow the rapid characterization of germplasm. This assists plant breeders in understanding the relationship between alleles and adaptation to specific environments and serves as a starting point for defining how many alleles of a given gene exist. Screening germplasm for new alleles, hitherto laborious, will become much more accessible with the development of new high throughput and low

cost sequencing methods. This, however, still leaves the problem of establishing what phenotypes are conferred by these novel alleles and hence which are best suited to use in individual applications. Detailed knowledge of how individual genes work should eventually permit a bioinformatic classification, or at least prioritization, of new alleles and this can be coupled with *in vitro* or model plant assays. A major problem with understanding the effects of different alleles will be the lack of a common genetic background. This is being addressed by crossing mutations into the barley variety 'Bowman' and these lines provide good material for the identification of linked markers and mapping (Dahleen et al. 2005). New mutation work for the development of barley TILLING populations is being undertaken in the varieties 'Optic' and 'Barke' (reviewed by Sreenivasulu et al. 2008) and 'Morex' (Talame et al. 2008). In these populations mutations can be directly linked to phenotype. In the longer term, replacing alleles using homologous recombination might become possible.

A highly significant feature of the *Vrn*, *Ppd* and *vrs1* genes is their recurrence in domestication events within and between species. This suggests two explanations which are not mutually exclusive. These genes may be particularly prone to mutation and so favourable alleles occur more frequently than in other genes that could potentially give similar phenotypes. This predicts that these genes will also have a higher frequency of neutral mutations (i.e. they would generally be more polymorphic) but there is little evidence to support this. Alternatively, the positions and roles these genes occupy in developmental pathways may make them accessible to selection. In other words their mutation can affect flowering time or inflorescence development without disrupting other aspects of growth and physiology, resulting in new forms of plants with superior agronomic performance. In either case, these genes are logical starting points for exploring new species and are priority targets for exploring the range of allelic diversity in germplasm collections.

The major *Vrn* and *Ppd* genes discussed above often appear in QTL studies but in addition there are many additional QTL whose effects on flowering are poorly understood. These include so called 'earliness *per se*' (*eps*) genes whose effects do not seem to depend on vernalization or photoperiod (Worland 1996). Flowering time QTL distinct from *Vrn* or *Ppd* genes have been reported on all chromosome groups in the Triticeae and to prioritize them for study it is useful to identify those that appear most frequently. An example of such a meta-QTL analysis in wheat is given by Hanocq et al. (2007). These additional and significant flowering time genes can be identified as cloning becomes easier and an interesting question is whether they will work by modulating the effects of the known major genes or whether they will act independently. A third possibility is proposed by Bonnin et al. (2008) for QTL on wheat chromosomes 7A and 7D. In this case flowering time is correlated with polymorphisms in the *FT* gene, suggesting that these QTL may be homoeoalleles of *Vrn-B3*, a known major gene on 7B (see Section 20.3.3).

20.7 Using Flowering and Inflorescence Genes in Triticeae Breeding

An understanding of the genes underlying the timing of flowering allows plant breeders to understand the basis of local adaptation and to tune plants for high productivity. This enables better use of exotic and wild germplasm as plants can rapidly be returned to locally adapted types. In addition, understanding how alleles act in combination allows a prediction of the adaptive limits of crops, and of local gene pools, which is significant for planning breeding responses to changes in climate. The ability to understand and manipulate inflorescence development will allow maximum grain fill in given situations, especially in cases where crops are currently sink limited (i.e. there is photosynthate that is not converted into grain yield). Other aspects such as changing morphology to assist hybrid production are also possible. In all these cases, the aim is to increase the ability of plant breeders to design crops that are well fitted to the needs of agronomy, whatever those needs might be.

Acknowledgments The author acknowledges support by the Biotechnology and Biological Sciences Research Council of Great Britain through grant-in-aid to the John Innes Centre.

References

- Achard, P., Cheng, H., De Grauwe, L., Decat, J., Schoutteten, H., Moritz, T., Van Der Straeten, D., Peng, J. and Harberd, N.P. (2006) Integration of plant responses to environmentally activated phytohormonal signals. *Science* 311, 91–94.
- Bäurle, I. and Dean, C. (2006) The timing of developmental transitions in plants. *Cell* 125, 655–664.
- Beales, J., Laurie, D.A. and Devos, K.M. (2005) Allelic variation at the linked *API* and *PhyC* loci in hexaploid wheat is associated but not perfectly correlated with vernalization response. *Theor. Appl. Genet.* 110, 1099–1107.
- Beales, J., Turner, A., Griffiths, S., Snape, J.W. and Laurie, D.A. (2007) A pseudo-response regulator is misexpressed in the photoperiod insensitive *Ppd-D1a* mutation of wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* 115, 721–733.
- Benlloch, R., Berbel, A., Serrano-Mislata, A. and Madueno, F. (2007) Floral initiation and inflorescence architecture: a comparative view. *Ann. Bot.* 100, 659–676.
- Blanc, G., Hokamp, K. and Wolfe, K.H. (2003) A recent polyploidy superimposed on older large-scale duplications in the *Arabidopsis* genome. *Genome Res.* 13, 137–144.
- Bommert, P., Satoh-Nagasawa, N., Jackson, D. and Hirano, H.Y. (2005) Genetics and evolution of inflorescence and flower development in grasses. *Plant Cell Physiol.* 46, 69–78.
- Bonnin, I., Rousset, M., Madur, D., Sourdille, P., Dupuits, C., Brunel, D. and Goldringer, I. (2008) *FT* genome A and D polymorphisms are associated with the variation of earliness components in hexaploid wheat. *Theor. Appl. Genet.* 116, 383–394.
- Chandler, P.M., Marion-Poll, A., Ellis, M. and Gubler, F. (2002) Mutants at the *Slender1* locus of barley cv Himalaya: molecular and physiological characterization. *Plant Physiol.* 129, 181–190.
- Chuck, G., Meeley, R.B. and Hake, S. (1998) The control of maize spikelet meristem fate by the *APETALA2*-like gene *indeterminate spikelet1*. *Genes Dev.* 12, 1145–1154.

- Chuck, G., Meeley, R., Irish, E., Sakai, H. and Hake, S. (2007) The maize *tasselseed4* microRNA controls sex determination and meristem cell fate by targeting *Tasselseed6/indeterminate spikelet1*. *Nat. Genet.* 39, 1517–1521.
- Clark, R.M., Wagler, T.N., Quijada, P. and Doebley, J. (2006) A distant upstream enhancer at the maize domestication gene *tb1* has pleiotropic effects on plant and inflorescent architecture. *Nat. Genet.* 38, 594–597.
- Cockram, J., Jones, H., Leigh, F.J., O'Sullivan, D., Powell, W., Laurie, D.A. and Greenland, A.J. (2007a) Control of flowering time in temperate cereals: genes, domestication and sustainable productivity. *J. Exp. Bot.* 58, 1231–1244.
- Cockram, J., Chiapparino, E., Taylor, S.A., Stamati, K., Donini, P., Laurie, D.A. and O'Sullivan, D.M. (2007b) Haplotype analysis of vernalization loci in European barley germplasm reveals novel *VRN-H1* alleles and a predominant winter *VRN-H1/VRN-H2* multi-locus haplotype. *Theor. Appl. Genet.* 115, 993–1001.
- Dahleen, L.S., Vander Wal, L.J. and Franckowiak, J.D. (2005) Characterization and molecular mapping of genes determining semidwarfism in barley. *J. Hered.* 96, 654–662.
- Danyluk, J., Kane, N.A., Breton, G., Limin, A.E., Fowler, D.B. and Sarhan, F. (2003) *TaVRT-1*, a putative transcription factor associated with vegetative to reproductive transition in cereals. *Plant Physiol.* 132, 1849–1860.
- Doi, K., Izawa, T., Fuse, T., Yamanouchi, U., Kubo, T., Shimatani, Z., Yano, M. and Yoshimura, A. (2004) *Ehd1*, a B-type response regulator in rice, confers short-day promotion of flowering and controls *FT*-like gene expression independently of *Hd1*. *Genes Dev.* 18, 926–936.
- Dubcovsky, J., Loukoianov, A., Fu, D., Valarik, M., Sanchez, A. and Yan, Y. (2006) Effect of photoperiod on the regulation of wheat vernalization genes *VRN1* and *VRN2*. *Plant Mol. Biol.* 60, 469–480.
- Dunford, R.P., Griffiths, S., Christodoulou, V. and Laurie, D.A. (2005) Characterisation of a barley (*Hordeum vulgare* L.) homologue of the *Arabidopsis* flowering time regulator *GIGANTEA*. *Theor. Appl. Genet.* 110, 925–931.
- Faris, J.D., Fellers, J.P., Brooks, S.A. and Gill, B.S. (2003) A bacterial artificial chromosome contig spanning the major domestication locus *Q* in wheat and identification of a candidate gene. *Genetics* 164, 311–321.
- Faure, S., Higgins, J., Turner, A. and Laurie, D.A. (2007) The *FLOWERING LOCUS T*-like gene family in barley (*Hordeum vulgare*). *Genetics* 176, 599–609.
- Fu, D., Szucs, P., Yan, L., Helguera, M., Skinner, J.S., von Zitzewitz, J., Hayes, P.M. and Dubcovsky, J. (2005) Large deletions within the first intron in *VRN-1* are associated with spring growth habit in barley and wheat. *Mol. Genet. Genomics* 273, 54–65.
- Fu, D., Dunbar, M. and Dubcovsky, J. (2007) Wheat *VIN3*-like PHD finger genes are up-regulated by vernalization. *Mol. Genet. Genomics* 277, 301–313.
- Fu, X.D. and Harberd, N.P. (2003) Auxin promotes *Arabidopsis* root growth by modulating gibberellin response. *Nature* 421, 740–743.
- Griffiths, S., Dunford, R.P., Coupland, G. and Laurie, D.A. (2003) The evolution of *CONSTANS*-like gene families in barley (*Hordeum vulgare*), rice (*Oryza sativa*) and *Arabidopsis thaliana*. *Plant Physiol.* 131, 1855–1867.
- Hanocq, E., Laperche, A., Jaminon, O., Lainé, A.L. and Le Gouis, J. (2007) Most significant genome regions involved in the control of earliness traits in bread wheat, as revealed by QTL meta-analysis. *Theor. Appl. Genet.* 114, 569–584.
- Hemming, M.N., Peacock, W.J., Dennis, E.S. and Trevaskis, B. (2008) Low-temperature and daylength cues are integrated to regulate *FLOWERING LOCUS T* in barley. *Plant Physiol.* 147, 355–366.
- Ikeda, A., Ueguchi-Tanaka, M., Sonoda, Y., Kitano, H., Koshioka, M., Futsuhara, Y., Matsuoka, M. and Yamaguchi, J. (2001) *Slender rice*, a constitutive gibberellin response mutant, is caused by a null mutation of the *SLR1* gene, an ortholog of the height-regulating gene *GAI/RGA/RHT/D8*. *Plant Cell* 13, 999–1010.

- Itoh, J., Sato, Y., Nagato, Y. and Matsuoka, M. (2006) Formation, maintenance and function of the shoot apical meristem in rice. *Plant Mol. Biol.* 60, 827–842.
- Iwaki, K., Nakagawa, K., Kuno, H. and Kato, K. (2000) Ecogeographical differentiation in east Asian wheat, revealed from the geographical variation of growth habit and *Vrn* genotype. *Euphytica* 111, 137–143.
- Kane, N.A., Danyluk, J., Tardif, G., Ouellet, F., Laliberte, J.F., Limin, A.E., Fowler, D.B. and Sarhan, F. (2005) *TaVRT-2*, a member of the StMADS-11 clade of flowering repressors, is regulated by vernalization and photoperiod in wheat. *Plant Physiol.* 138, 2354–2363.
- Kane, N.A., Agharbaoui, Z., Diallo, A.O., Adam, H., Tominaga, Y., Ouellet, F. and Sarhan, F. (2007) *TaVRT2* represses transcription of the wheat vernalization gene *TaVRN1*. *Plant J.* [Epub ahead of print].
- Kellogg, E.A. (2007) Floral displays: genetic control of grass inflorescences. *Curr. Opin. Plant Biol.* 10, 26–31.
- Kobayashi, Y. and Weigel, D. (2007) Move on up, it's time for change – mobile signals controlling photoperiod-dependent flowering. *Genes Dev.* 21, 2371–2384.
- Kojima, S., Takahashi, Y., Kobayashi, Y., Monna, L., Sasaki, T., Araki, T. and Yano, M. (2002) *Hd3a*, a rice ortholog of the Arabidopsis *FT* gene, promotes transition to flowering downstream of *Hd1* under short-day conditions. *Plant Cell Physiol.* 43, 1096–1105.
- Komatsuda, T., Pourkheirandish, M., He, C., Azhaguvel, P., Kanamori, H., Perovic, D., Stein, N., Graner, A., Wicker, T., Tagiri, A., Lundqvist, U., Fujimura, T., Matsuoka, M., Matsumoto, T. and Yano, M. (2007) Six-rowed barley originated from a mutation in a homeodomain-leucine zipper I-class homeobox gene. *Proc. Natl. Acad. Sci. USA* 104, 1424–1429.
- Law, C.N. and Worland, A.J. (1997) Genetic analysis of some flowering time and adaptive traits in wheat. *New Phytol.* 137, 19–28.
- Loukoianov, A., Yan, L.L., Blechl, A., Sanchez, A. and Dubcovsky, J. (2005) Regulation of *VRN-1* vernalization genes in normal and transgenic polyploid wheat. *Plant Physiol.* 138, 2364–2373.
- Malcomber, S.T., Preston, J.C., Reinheimer, R., Kossuth, J. and Kellogg, E.A. (2006) Developmental gene evolution and the origin of grass inflorescence diversity. *Adv. Bot. Res.* 44, 425–481.
- Mizuno, T. and Nakamichi, N. (2005) Pseudo-response regulators (*PRRs*) or True oscillator components (*TOCs*). *Plant Cell Physiol.* 46, 677–685.
- Moore, G., Gale, M.D., Kurata, N. and Flavell, R.B. (1993) Molecular analysis of small grain cereal genomes: current status and prospects. *Bio Technol.* 11, 584–589.
- Moore, G., Devos, K.M., Wang, Z. and Gale, M.D. (1995) Grasses, line up and form a circle. *Curr. Biol.* 5, 737–739.
- Müller, K.J., Romano, N., Gerstner, O., Garcia-Marato, F., Pozzi, C., Salamini, F. and Rohde, W. (1995) The barley *Hooded* mutation caused by a duplication in a homeobox gene intron. *Nature* 374, 727–730.
- Murai, K., Miyamae, M., Kato, H., Takumi, S. and Ogihara, Y. (2003) *WAP1*, a wheat *APETALA1* homolog, plays a central role in the phase transition from vegetative to reproductive growth. *Plant Cell Physiol.* 44, 1255–1265.
- Peng, J., Carol, P., Richards, D.E., King, K.E., Cowling, R.J., Murphy, G.P. and Harberd, N. P. (1997) The Arabidopsis *GAI* gene defines a signalling pathway that negatively regulates gibberellin responses. *Genes Dev.* 11, 3194–3205.
- Peng, J.R., Richards, D.E., Hartley, N.M., Murphy, G.P., Devos, K.M., Flintham, J.E., Beales, J., Fish, L.J., Worland, A.J., Pelica, F., Sudhakar, D., Christou, P., Snape, J.W., Gale, M.D. and Harberd, N.P. (1999) 'Green revolution' genes encode mutant gibberellin response modulators. *Nature* 400, 256–261.
- Pourkheirandish, M., Wicker, T., Stein, N., Fujimura, T. and Komatsuda, T. (2007) Analysis of the barley chromosome 2 region containing the six-rowed spike gene *vrsl* reveals a

- breakdown of rice-barley micro collinearity by a transposition. *Theor. Appl. Genet.* 114, 1357–1365.
- Preston, J.C. and Kellogg, E.A. (2007) Conservation and divergence of *APETALAI/FRUITFULL*-like gene function in grasses: evidence from gene expression analyses. *Plant J.* 52, 69–81.
- Preston, J.C. and Kellogg, E.A. (2008) Discrete developmental roles for temperate cereal grass *VRN1/FUL*-like genes in flowering competency and the transition to flowering. *Plant Physiol.* 10.1104/pp.107.109561.
- Rossini, L., Vecchietti, A., Shicolas, L., Stein, N., Franzago, S., Salamini, F. and Pozzi, C. (2006) Candidate genes for barley mutants involved in plant architecture: an *in silico* approach. *Theor. Appl. Genet.* 112, 1073–1085.
- Salvi, S., Sponza, G., Morgante, M., Tomes, D., Niu, X., Fengler, K.A., Meeley, R., Ananiev, E.V., Svitashov, S., Bruggemann, E., Li, B., Hainey, C.F., Radovic, S., Zaina, G., Rafalski, J.A., Tingey, S.V., Miao, G.H., Phillips, R.L. and Tuberosa, R. (2007) Conserved noncoding genomic sequences associated with a flowering-time quantitative trait locus in maize. *Proc. Natl. Acad. Sci. USA* 104, 11376–11381.
- Satoh-Nagasawa, N., Nagasawa, N., Malcomber, S., Sakai, H. and Jackson, D. (2006) A trehalose metabolic enzyme controls inflorescence architecture in maize. *Nature* 441, 227–230.
- Schmitz, R.J. and Amasino, R.M. (2007) Vernalization: a model for investigating epigenetics and eukaryotic gene regulation in plants. *Biochim. Biophys. Acta* 1769, 269–275.
- Shindo, C., Aranzana, M.J., Lister, C., Baxter, C., Nicholls, C., Nordborg, M. and Dean, C. (2005) Role of *FRIGIDA* and *FLOWERING LOCUS C* in determining variation in flowering time of *Arabidopsis*. *Plant Physiol.* 138, 1163–1173.
- Shindo, C., Lister, C., Crevillen, P., Nordborg, M. and Dean, C. (2006) Variation in the epigenetic silencing of *FLC* contributes to natural variation in *Arabidopsis* vernalization response. *Genes Dev.* 20, 3079–3083.
- Shitsukawa, N., Ikari, C., Shimada, S., Kitagawa, S., Sakamoto, K., Saito, H., Ryuto, H., Fukunishi, N., Abe, T., Takumi, S., Nasuda, S. and Murai, K. (2007) The einkorn wheat (*Triticum monococcum*) mutant, *maintained vegetative phase*, is caused by a deletion in the *VRN1* gene. *Genes Genet. Syst.* 82, 167–170.
- Simons, K.J., Fellers, J.P., Trick, H.N., Zhang, Z., Tai, Y.-S., Gill, B.S. and Faris, J.D. (2006) Molecular characterization of the major wheat domestication gene *Q*. *Genetics* 172, 547–555.
- Slade, A.J., Fuerstenberg, S.I., Loeffler, D., Steine, M.N. and Facciotti, D. (2005) A reverse genetic, nontransgenic approach to wheat crop improvement by TILLING. *Nat. Biotechnol.* 23, 75–81.
- Sreenivasulu, N., Graner, A. and Wobus, U. (2008) Barley genomics: an overview. *Int. J. Plant Genomics* doi: 10.1155/2008/486258.
- Sung, S.B. and Amasino, R.M. (2004) Vernalization in *Arabidopsis thaliana* is mediated by the PHD finger protein *VIN3*. *Nature* 427, 159–164.
- Talame, V., Bovina, R., Sanguineti, M.C., Tuberosa, R., Lundqvist, U. and Salvi, S. (2008) TILLMore, a resource for the discovery of chemically induced mutants in barley. *Plant Biotechnol. J.* 5, 477–485.
- Trevaskis, B., Hemming, M.N., Peacock, W.J. and Dennis, E.S. (2006) *HvVRN2* responds to daylength, whereas *HvVRN1* is regulated by vernalization and developmental status. *Plant Physiol.* 140, 1397–1405.
- Trevaskis, B., Hemming, M.N., Dennis, E.S. and Peacock, W.J. (2007) The molecular basis of vernalization-induced flowering in cereals. *Trends Plant Sci.* 12, 352–357.
- Turner, A., Beales, J., Faure, S., Dunford, R.P. and Laurie, D.A. (2005) The pseudo-response regulator *Ppd-H1* provides adaptation to photoperiod in barley. *Science* 310, 1031–1034.
- Vollbrecht, E., Veit, B., Sinha, N. and Hake, S. (1991) The developmental gene *Knotted-1* is a member of a maize homeobox gene family. *Nature* 350, 241–243.

- von Zitzewitz, J., Szucs, P., Dubcovsky, J., Yan, L., Francia, E., Pecchioni, N., Casas, A., Chen, T.H., Hayes, P.M. and Skinner, J.S. (2005) Molecular and structural characterization of barley vernalization genes. *Plant Mol. Biol.* 59, 449–467.
- Winichayakul, S., Beswick, N.L., Dean, C. and Macknight, R.C. (2005) Components of the *Arabidopsis* autonomous floral promotion pathway, *FCA* and *FY*, are conserved in monocots. *Funct. Plant Biol.* 32, 345–355.
- Worland, A.J. (1996) The influence of flowering time genes on environmental adaptability in European wheats. *Euphytica* 89, 49–57.
- Yan, L., Loukoianov, A., Tranquilli, G., Helguera, M., Fahima, T. and Dubcovsky, J. (2003) Positional cloning of the wheat vernalization gene *VRN1*. *Proc. Natl. Acad. Sci. USA* 100, 6263–6268.
- Yan, L., Loukoianov, A., Blechl, A., Tranquilli, G., Ramakrishna, W., SanMiguel, P., Bennetzen, J.L., Echenique, V. and Dubcovsky, J. (2004a) The wheat *VRN2* gene is a flowering repressor down-regulated by vernalization. *Science* 303, 1640–1644.
- Yan, L., Helguera, M., Kato, K., Fukuyama, S., Sherman, J. and Dubcovsky, J. (2004b) Allelic variation at the *VRN-1* promoter region in polyploid wheat. *Theor. Appl. Genet.* 109, 1677–1686.
- Yan, L., Fu, D., Li, C., Blechl, A., Tranquilli, G., Bonafede, M., Sanchez, A., Valarik, M., Yasuda, S. and Dubcovsky, J. (2006) The wheat and barley vernalization gene *VRN3* is an orthologue of *FT*. *Proc. Natl. Acad. Sci. USA* 103, 19581–19586.
- Yu, J., Wang, J., Lin, W., Li, S., Li, H., Zhou, J., Ni, P., Dong, W. et al. (2005) The genomes of *Oryza sativa*: a history of duplications. *PLoS Biol.* 3, e38 (0266–0281).

Chapter 21

Genomics of Quality Traits

W. Ma, O. Anderson, H. Kuchel, Y. Bonnardeaux, H. Collins, M.K. Morell, P. Langridge, and R. Appels

Abstract The quality attributes of cereal grains are valued in the context of a complex food chain that integrates outputs achievable by breeding, production, and processing. New processing technologies, environmental change, and changes in consumer preferences demand that quality attributes of wheat and barley need to be continually modified. The advances in the genomics of quality described in this chapter provide the basis for ensuring that the genetic approaches encompassing the complexities of the gene networks underpinning quality attributes can meet the challenges presented by the rapid changes occurring within the food chain.

21.1 Introduction

The wheat and barley industry is currently undergoing major changes as the value of the grain increases due to the combined pressures of limited world-stocks resulting from droughts and loss of available land, and competition for use of the grain for biofuels and feed. The marketers and growers of major crops such as wheat and barley need to increase their productivity by 10–20% over the next decade (utilizing a reduced growing area) while maintaining the quality attributes demanded by increasingly sophisticated customers. The market in the India-Asia region is growing most rapidly and wheat/barley quality improvement has a significant part to play in developing the linkages with customers in these markets. New advances in the engineering of large-scale processing equipment as well as changes in the market place resulting from the striking demographic changes in the region indicate that traditional quality attributes of wheat and barley will also need to evolve. Genome level studies are considered crucial to defining the genes and gene networks underpinning the functional qualities of flour and malt for a range of end products and health claims.

W. Ma (✉)

Department of Agriculture and Food, Perth, Western Australia
e-mail: w.ma@murdoch.edu.au

The grain of the cereal crops provides a large portion of both the total calories and protein in the human diet. Many of the chromosome regions affecting quality attributes are now amenable to analysis at the genomic structure/function level and in this chapter the aspects contributing to wheat and barley quality attributes targeted to specific end products are discussed. Although we have described some recent work on barley cell wall structure and the implications for this work on human health, the broader aspects of nutrition for wheat and barley are not within the scope of this chapter.

21.2 Genomics of Barley Quality

Barley grain is used for human food, animal feed and for the production of malt. Malt finds its way into many food products but the most important are beer and whiskey.

Animal feed qualities for barley are of interest as the cost of grain increases due to competition from the biofuels area. Although the emphasis for feed use is on total yield, the requirements for monogastric animals (pigs and poultry) are quite different from those for ruminants (feedlot and dairy), (Gali et al. 1998; White et al. 2007). Pigs and poultry demand a high available energy intake, with easily digestible starch, while for ruminants the industry requires grains that are slower fermenting to avoid acidosis. Grain as a source of carbohydrate is a key focus.

Traditionally, feed barley has been spring or winter type and six rowed and was selected primarily for yield while malting barley has been exclusively spring type and in most cases two rowed. Wild barley (*Hordeum vulgare* ssp *spontaneum*) is also two rowed but a mutation at a single gene *Vrs1* led to the six row forms which were generally higher yielding. A positional cloning approach was recently used to isolate the *Vrs1* gene (Komatsuda et al. 2007). The six rowed phenotype resulted from loss of function of *Vrs1* and conversion of rudimentary lateral spikelets into fully functional spikelets.

A second gene that has been important for both feed and human food barley is the hulless locus (*nud*) on 7H. In hulless lines the glumes detach from the pericarp during threshing to produce a hulless or naked grain. Several countries have now established breeding programs focused on producing hulless or naked barley specifically for the animal feed industries. The gene underlying this trait was recently cloned and encodes an ethylene responsive transcription factor (Taketa et al. 2008).

21.2.1 Human Food

Currently barley is largely consumed by poor farmers in Northern Africa and Asia. The highest levels of barley consumption are in Morocco where farmers consume on average 54 Kg per annum. In the highlands of Ethiopia barley

accounts for over 60% of the food consumed. It is also grown extensively for human food in the Himalayas, particularly Nepal, in the mountainous regions of Yemen and in the Andes, notably in Columbia, Peru, Bolivia and Ecuador (Grando 2002).

Recently, barley has attracted increased attention in the developed world due to its relatively high content of non-starch carbohydrates. Different classes of non-starch carbohydrates have been found to stimulate the immune system, to reduce cholesterol levels and to be anti-tumourogenic (Mantovani et al. 2008). In barley, the levels of (1,3;1,4)- β -D-glucan in cell walls of the starchy endosperm can be 70% or more by weight. While barley, oat, and rye grains are all rich sources of (1,3;1,4)- β -D-glucan, wheat, rice, and maize have much lower concentrations. The genes encoding the enzymes involved in (1,3;1,4)- β -D-glucan synthesis were recently cloned from barley (Burton et al. 2006). A combination of genetic information and candidate gene analysis was used to isolate this group of genes. The chemical similarities between cellulose and (1,3;1,4)- β -D-glucans suggested that the (1,3;1,4)- β -D-glucan synthases might be members of one of the Cellulose synthase like (*Cs/*) gene families (Fig. 21.1). In most plants the

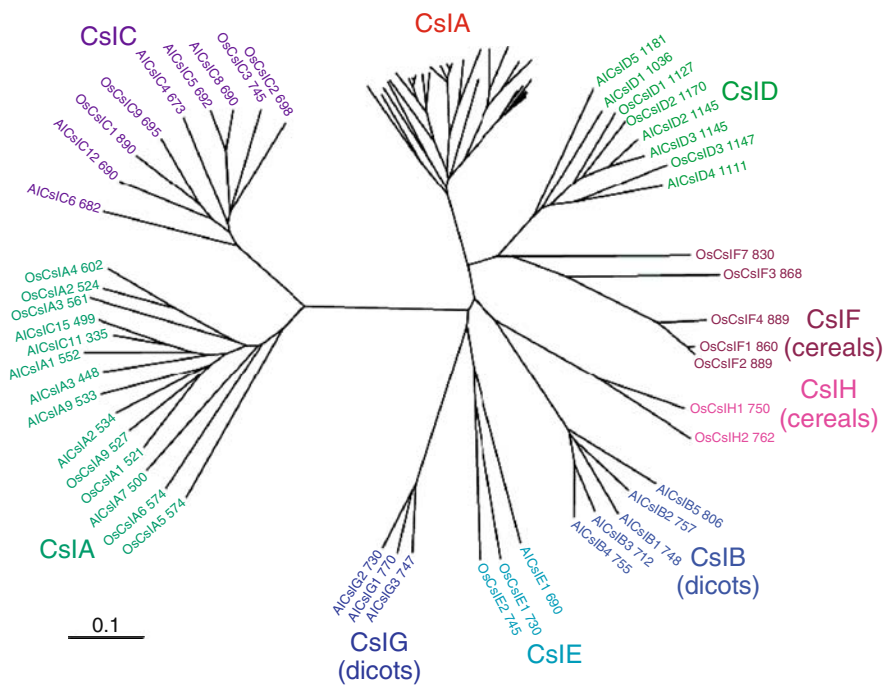


Fig. 21.1 The cellulose synthase (*CesA*) and cellulose synthase-like (*Csl*) gene families from higher plants contain around 50 members. The *CesA* genes are involved in cellulose biosynthesis but the functions of most of the *Csl* genes have not been defined (modified from Lerouxel et al. 2006) (See Color Insert)

Csl gene families are large and are divided into sub-groups, named *CslA* to *CslH*. Two groups, *CslF* and *CslH*, are only found in the cereals. Given that (1,3;1,4)- β -D-glucan is also only found in cereals, these two gene families appeared to be likely candidates for the synthases. By using comparative genomics and the known position of a major QTL affecting barley (1,3;1,4)- β -D-glucan levels, Burton et al. (2006) identified the *CslF* genes as the most likely candidates for (1,3;1,4)- β -D-glucan synthases. When the *CslF* genes were expressed in *Arabidopsis*, (1,3;1,4)- β -D-glucan was detected in walls of transgenic plants (Burton et al. 2006) confirming that the genes participated in the (1,3;1,4)- β -D-glucan synthesis.

The cloning of these genes offers the potential to modify the levels of (1,3;1,4)- β -D-glucan in barley and other cereal grains. High levels of (1,3;1,4)- β -D-glucan in wheat would be beneficial to human health, where they represent soluble dietary fibre and appear to reduce the risks of colorectal cancer, high serum cholesterol and cardiovascular disease, obesity and non-insulin dependent diabetes (Mantovani et al. 2008). Conversely, (1,3;1,4)- β -D-glucan is an anti-nutritional factor for monogastric animals and reduced levels would be desirable. Reduced levels are also desired by the brewing industry where (1,3;1,4)- β -D-glucan can cause problems in beer filtration. There may also be opportunities to use lowered levels of (1,3;1,4)- β -D-glucan in the biofuels industry. Therefore, the cloning of the genes involved in (1,3;1,4)- β -D-glucan synthesis provides the basis for the production of novel grains via genetic engineering.

21.2.2 Malting and Brewing

In barley, malting quality has been well studied and improved profiles have included whole grain qualities (colour and plumpness), rate of germination, specific enzyme levels to match the malting process, and thermo-stability of key enzymes.

A large number of components of malt quality are routinely screened in the quality assessment of malting barley. The most important are listed in Table 21.1 (see also Hamilton and Lewis 1974).

Malting and brewing essentially involves the enzymatic degradation of the barley endosperm and the fermentation of the released sugars. The first stage is the hydrolysis of the crushed cell layer surrounding the endosperm and of the cell walls within the endosperm. With the high content of (1,3;1,4)- β -D-glucan in the endosperm cell walls (1,3;1,4)- β -D-glucanase is of prime importance. There are two enzymes EI and EII that are responsible for this process. Proteases also play an active role in facilitating hydrolysis of the starch granules but are also involved in activation of other malt enzymes including protein Z, α -amylase and β -amylase. A recent transcript analysis that compared malt and feed barley identified carboxypeptidase as a key determinant of malt quality (Potokina et al. 2006).

Table 21.1 A summary of the major characteristics used to assess malt quality (from Eglinton 2003)

Malt quality parameters	Definition
Moisture (%)	Percentage of water in the malt
Extract, Fine (%)	Level of water soluble material extracted from the malt after different levels of milling
Extract, Coarse (%)	
Fine/Coarse Extract Difference (%)	Provides a simple index of modification
Wort colour	The intensity of the colour of the wort, reflecting the internal colour of the grain
Diastatic Power (DP)	Total activity of the enzymes involved in starch hydrolysis
Alpha Amylase	α -amylase enzyme activity extracted from malt
Beta Amylase	β -amylase enzyme activity extracted from malt
Fermentability (AAL%)	The level of fermentable sugars obtained from the malt
Total Nitrogen	Total nitrogen content of the malt, this is also a measure of total protein
Soluble Nitrogen	Total nitrogen content of the malt, this is also a measure of the extent of modification
Kolbach Index	Ratio of soluble N to total N, providing a measure of the extent of modification
Wort Beta Glucan	Total (1-3),(1-4) beta glucan in the wort
Wort Viscosity	Reflects the level of beta glucan and other soluble high molecular weight material

Proteinase inhibitors also play a role in controlling the rate of endosperm digestion. In particular, protein Z (a serine proteinase) has been implicated in influencing endosperm hydrolysis (Hejgaard et al. 1985).

21.2.2.1 β -amylase

Several enzymes are involved in the hydrolysis of the starch. These include α -amylase, limit dextrinase, α -glucosidase and β -amylase. The genes encoding these enzymes have all been cloned and several groups have investigated natural variation in the properties of these enzymes and the corresponding genes. However, in QTL studies β -amylase is the only enzyme that consistently shows an association with diastatic power, a key determinant of malt quality (see below). There is also a strong association between β -amylase activity and wort attenuation (the amount of un-degraded starch that remains due to incomplete hydrolysis).

There are two forms of β -amylase in barley. One is endosperm specific, accumulates during seed development and is encoded by *Bmy1* on the long arm of chromosome 4H. This enzyme is a major component of grain protein. The second enzyme is not as tissue specific, is encoded by *Bmy2* locus on chromosome 2H, and does not play a significant role in the malting processes. In most commercial malts β -amylase activity appears to be limiting although the efficiency of this enzyme is influenced by other enzymes involved in starch

degradation, particularly limit dextrinase (MacGregor et al. 1999). This issue is particularly important where barley malt is used in conjunction with adjuncts, usually derived from rice or maize. The quantity of β -amylase *per se* does not seem to be the major concern since this protein makes up 1–2% of the total barley protein in malt (MacGregor et al. 1971). It is more probable that variation in the thermostability and substrate affinity of β -amylase are responsible for the variation seen in diastatic power. During brewing, the malt passes through a mashing phase where the temperature can reach 70°C. β -amylase has maximal activity at 45°C but its activity starts to decline as the temperature goes above 55°C (Yoshigi et al. 1995). Three β -amylase alleles, *Bmy1*-Sd1, *Bmy1*-Sd2L and *Bmy1*-Sd2H are found in commercial barley varieties although several further alleles have been found in wild barleys. These have medium, low and high thermostability, respectively (Eglinton et al. 1998).

The various β -amylase alleles have been cloned and sequenced and active β -amylase enzyme can be readily produced in bacterial (*E. coli*) expression systems. Consequently, it has been possible to use site-directed mutagenesis to characterize the effects of each amino acid variant on the activity and thermostability of the enzyme (Ma et al. 2001). These results indicated that the variation in thermostability was due to two amino acid substitutions which changed the thermostability by around 2°C each (Ma et al. 2001). The variation in activity was also due to single amino acid differences. Based on these findings, Ma and colleagues (2001) generated a novel β -amylase gene that combined optimal specific activity and thermostability. The detailed structural information now available on the β -amylase enzyme and the ability to engineer new variants of this enzyme with enhanced properties for the brewing industry offers new options for genetic engineering of barley (Ma et al. 2000). While these options have been available for several years, they are yet to be translated to commercial reality. However, they have allowed researchers to develop sophisticated strategies for screening wild germplasm for new “natural” variants (Eglinton 2003).

21.2.3 QTL associated with malting quality

Hayes et al. (2001) reported that there are 181 QTL described in the literature for 29 barley and malt quality phenotypes. These are summarized on the web site <http://www.css.orst.edu/barley/nabgmp/qtlsum.htm>. This summary shows a BinMap created by Kleinhofs and Han (2002). Each chromosome has been divided into a number of bins and the QTL have been assigned to these bins. QTL for malt extract have been identified in 8 populations developed by researchers from all over the world. These are schematically shown in Fig. 21.2, which is based on this BinMap.

The most extensively studied population is Steptoe/Morex. Malt extract was first measured in this population in 1991 from four sites and this was repeated in 1992 at a further five sites (Hayes et al. 1993; Hayes and Iyamabo 1994;

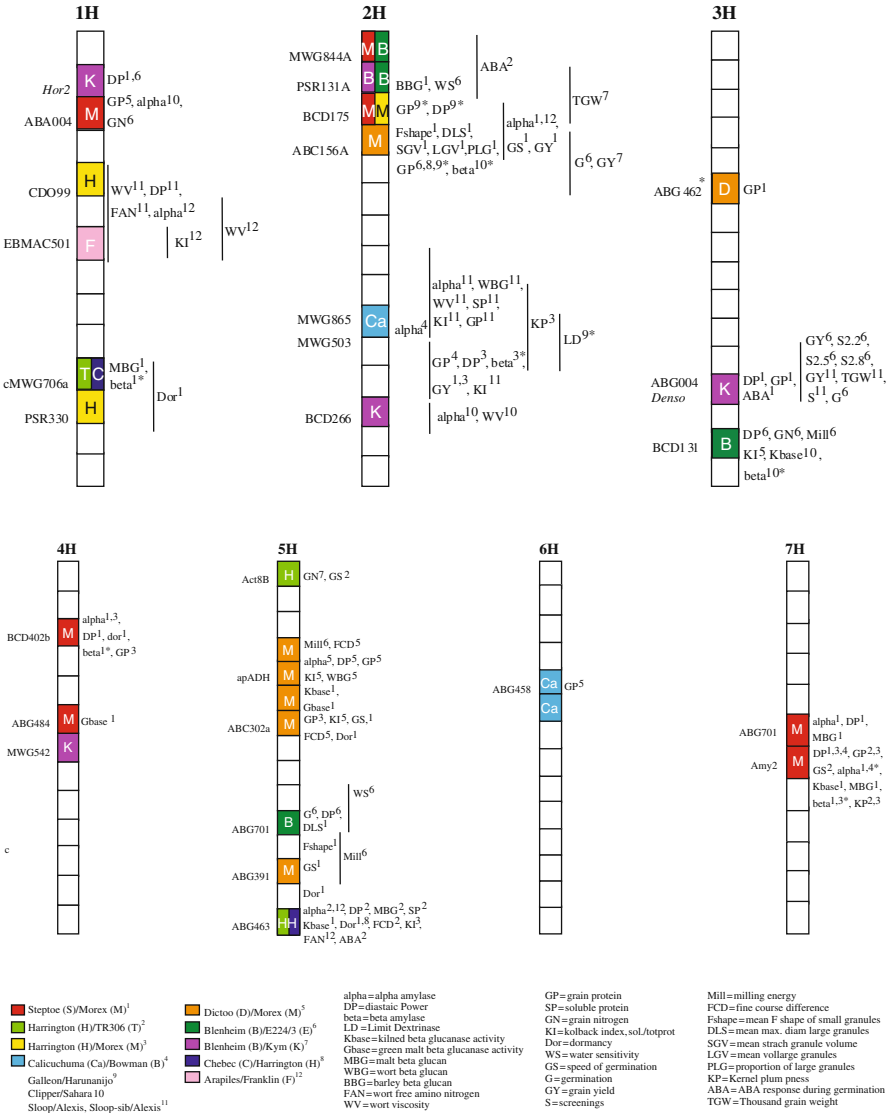


Fig. 21.2 A schematic representation of malt extract QTL, based on Hayes et al. (1997, 2001), <http://www.css.orst.edu/barley/nabgmp/qtlsum.htm>. Each chromosome is separated into a number of regions called Bins as described by Kleinhofs and Han (2002). Colored squares represent regions found in mapping populations around the world and indicate the locations of the malt extract QTL locations – the letters in the colored boxes indicate the first letter of the parent contributing the malt extract QTL allele. Markers are listed left of each chromosome. Listed right of the chromosome are other traits found to be associated with each region. As discussed in the text the populations analyzed to produce the distribution of QTL include (1) *Stepoe/Morex*: Two “six row” varieties, grown and mapped in USA. *Stepoe* is a low extract feed quality variety and *Morex* is a malting quality variety. (2) *Dictoo/Morex*: Two “six row” varieties. Also grown and mapped in the USA. *Dictoo* is a winter variety and *Morex*

Ullrich et al. 1997). QTL for malt extract were identified on all chromosomes except 3H. However, a number of these were only identified using data from individual sites and consequently have not been included in the summary by Hayes et al. (2001). All of the QTL in this summary, from this population, have Morex donating the higher malt extract allele (Fig. 21.1). However, Steptoe is responsible for two of the QTL that were identified using data from a single site only (Hayes and Iyamabo 1994).

Morex was also used as a parent to develop two other mapping populations, namely Harrington/Morex and Dictoo/Morex. Only one region, on the short arm of chromosome 2H, was found to be significantly associated with malt extract in all three of these populations. Morex is responsible for donating the higher malt extract allele in this region, for all three populations. Four other significant regions were also found to be associated with malt extract in the Dictoo/Morex population, three regions on chromosome 5H and one on chromosome 3H (Oziel et al. 1996).

Two other significant regions were found to be associated with malt extract in the Harrington/Morex mapping population, both on chromosome 1H (Table 21.1) (Marquez-Cedillo et al. 2000). Harrington is responsible for donating the high extract allele at these two regions. The region on the long arm of chromosome 1H is flanked by a region found to be associated with malt extract in two other populations with Harrington as a parent (Harrington/TR306 and Chebec/Harrington). However in both of these cases Harrington donated the low extract allele (Mather et al. 1997; Hayes et al. 2001; Collins et al. 2003).

A region was found to be associated with malt extract on the long arm of chromosome 5H in the two populations Chebec/Harrington and Harrington/R306 (Mather et al. 1997; Hayes et al. 2001; Collins et al. 2003). Harrington donated the high extract allele in both populations. Malt extract was also found to be associated with a region on the short arm of chromosome 5H, in the population Harrington/TR306. This region was not found to be associated with malt extract in any other population.



Fig. 21.2 (continued) is a spring variety. (3) *Harrington/TR306*: Two “two row” varieties. This population was grown and mapped in Canada. Harrington is a high extract malting variety and TR306 is a feed quality line. (4) *Harrington/Morex*: A cross between two high extract malting quality varieties. These two varieties are the “two row” and “six row” malting quality standards for North America. (5) *Calicuchuma-sib/Bowman*: Calicuchuma-sib is an ICARDA/CIMMYT “six row” variety; Bowman is a “two row” variety. This population was grown and mapped in the USA. (6) *Blenheim/E224/3* and *Blenheim/Kym*: These populations were grown in the UK. They are the only populations where the extracts were measured using the IOB method. (7) *Chebec/Harrington*: This was grown and mapped in Australia by the National Barley Molecular Marker Program (NBMMP; Barr et al. 2003; Pallotta et al. 2003). (8) Two other populations, developed in Australia, have been used for mapping malting quality traits but not malt extract. These are Galleon/Haruna Nijo and Clipper/Sahara (Karakousis et al. 2003a,b) (See Color Insert)

The other two populations with a common parent are Blenheim/E224/3 and Blenheim/Kym (Thomas et al. 1996; Bezant et al. 1997; Powell et al. 1997). Thomas et al. (1996) found 18 regions in the Blenheim/E224/3 mapping population that were associated with malt extract, and another 22 regions that were associated with malt extract when adjusted to a grain nitrogen level of 1.5%. Of these, only four regions were found to be associated with malt extract at more than one site. Powell et al. (1997) found a further three regions associated with malt extract in that population, one of which was in common with the regions Thomas et al. (1996) found. Three of these regions could not be assigned to a bin and are not shown in Fig. 21.2. Although the variety Kym generally has a lower malt extract than Blenheim, it was responsible for donating the higher-level allele at five of the eight regions found to be associated with malt extract in the Blenheim/ Kym population (Bezant et al. 1997). Three regions could not be assigned to a bin and are not shown in Fig. 21.2. Only a single region was found to be significantly associated with HWE in the Blenheim/E224/3 and Blenheim/ Kym populations (Thomas et al. 1996; Bezant et al. 1997; Powell et al. 1997). This region is on the short arm of chromosome 2H and has Blenheim donating the higher malt extract allele.

QTL have been found for many other traits associated with malting quality (see Fig. 21.2). A single region of the genome is often found to influence a number of different traits. The terminal chromosome 5HL region, for example, carries major genetic factors for seed dormancy, pre-harvest sprouting, high malt extract, diastatic power, alpha amylase, wort free amino nitrogen, soluble protein and reduced levels of malt β -glucan and fine course difference. Dissecting the details of this region at the genomic level is of particular interest. Whether this region contains a gene “cluster”, number of closely linked individual genes controlling each trait separately or a single gene that has pleiotropic effects on each trait is currently unknown. A particular gene of interest located in this region is GA 20-oxidase, as discussed in Section 21.2.4.

In total more than 24 individual chromosome regions have been found to be associated with malt extract (Fig. 21.2) and the use of molecular markers linked to these regions can improve the effectiveness of selection of new varieties in barley breeding programs.

21.2.4 Germination as a Key Variable in Barley Quality

Susceptibility to preharvest sprouting of the grain while it is still in the head has become a significant challenge since the dormancy required to combat this susceptibility is detrimental to the germination required for the malting process. The challenge for breeders is to develop a barley genotype that is dormant enough to withstand preharvest sprouting but not too deeply dormant so that the grain does not germinate during the malting. Generally, there are two main phenotypes in the expression of dormancy, those that release dormancy

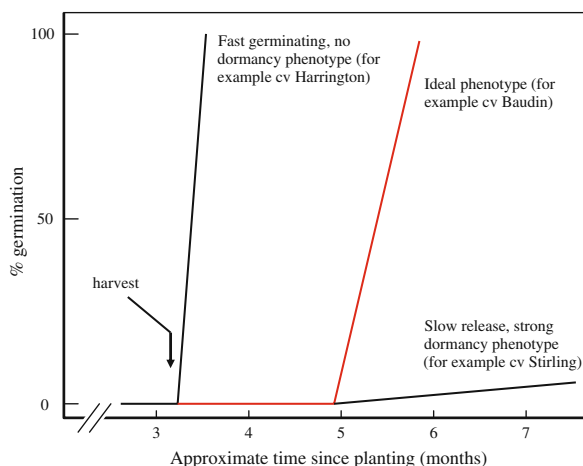


Fig. 21.3 Hypothetical graph demonstrating the germination behaviour of a fast germination phenotype with no dormancy such as the Harrington cultivar, a slow release phenotype with strong levels of seed dormancy such as Stirling and what may be the ideal seed dormancy phenotype displaying strong dormancy with a sudden timely release of dormancy such as Baudin (*See Color Insert*)

gradually and those that remain dormant for several months with a sudden release in dormancy. Ideally, a barley malting cultivar would be deeply dormant (primary dormancy) for the months during grain filling with a sudden release of dormancy immediately post harvest to reduce the cost of storage time (Fig. 21.3).

The QTL analyses performed on the Stirling/Harrington population in Bonnardeaux et al. (2008) identified two seed dormancy QTL on the 5H chromosome, that were consistent with the SD1 and SD2 QTL detected in several seed dormancy studies using feed varieties (Gao et al. 2003; Han et al. 1996, 1999; Oberthur et al. 1995). The SD1 and SD2 loci have been proposed to have different physiological functions, with SD1 affecting the formation and maintenance of dormancy and SD2 associated with the release of dormancy (Prada et al. 2004). The near-centromeric QTL on 5H (SD1) has been reported as the major QTL in previous studies and the QTL on the long arm of 5H (SD2) as a minor QTL. In some studies, however, the major QTL is on the long arm (SD2, Bonnardeaux et al. 2008). Therefore, the SD2 QTL may be more relevant to malting varieties and Australian germplasm. In a recent study Lin et al. (manuscript in press), mapped the SD1 and SD2 QTL for seed dormancy from a cross of Morex/Harrington and in contrast to the non-dormancy allele normally being derived from Morex (Han et al. 1996) the Morex allele was the dormant allele in this population. Thus, the nature of the QTL for seed dormancy on chromosome 5H needs further study. Li et al. (2004) conducted a comparative genomics study using rice, wheat and barley in the region of the SD2 locus on the long arm of 5H and suggested the gibberellin (GA) 20-oxidase

(encoded by the GA5 gene in *Arabidopsis*) was the underlying gene for this QTL. Since then, further evidence implicating the role of GA 20-oxidase in seed dormancy has emerged. Appleford et al. (2006) observed high levels of expression of the enzyme as well as GA 3-oxidase, a member of the same gene family, in developing and germinating wheat grains. In addition the GA 20-oxidase homologues are located in map positions in wheat chromosomes 5BL, 5DL and 4AL, syntenic to the region of the 5HL QTL in barley and consistent with an involvement in seed dormancy (Li et al. 2004; Appleford et al. 2006). The GA 20-oxidase is therefore a clear candidate gene for the control of dormancy. Another candidate gene is the *Viviparous1* gene (*Vp1*, Hobo et al. 1999), equivalent to the *ABI3* transcription factor in *Arabidopsis*, and is located within the dormancy QTL region on chromosome 3 of wheat (Bailey et al. 1999; McKibbin et al. 2002) which is also found on 3H of barley. The *Vp1* product is involved in the regulation of ABA signaling (Hattori et al. 1992; Susuki et al. 2003; reviewed in Finkelstein et al. 2008).

Environmental factors, such as moisture in particular, could be expected to have their greatest influence following the release of primary dormancy and thus affecting after-ripening or when the seed is cycling between secondary dormancy and a quiescent state. Germination of seeds in these phases predominantly relies on favorable environmental conditions. Minor QTL or those detected only in certain environments may represent genes that are more responsive to the environment and become active in specific environmental conditions. Lee et al. (2002) characterized the first gene connecting environmental and endogenous interactions in seed germination. The *Arabidopsis* RGL2 gene encoding a transcription factor, is induced by moisture and commences signaling by GA (Lee et al. 2002). Identification of genes such as the RGL2, that act at the interface between environmental cues and hormone response would elucidate the pathways by which environmental factors influence variation in germination (reviewed in Finkelstein et al. 2008).

21.3 Genomics of Wheat Quality

Biochemical and genetic studies have indicated that the properties of a group of well defined wheat storage proteins largely determine the dough rheological properties (Wall 1979; Branlard and Dardevet 1985; reviewed in Gras et al. 2001; Howitt et al. 2003) of wheat flour, a key quality attribute. Modification of quality can occur in a predictable way based on the genetic complement or in an unpredictable way as a result of environmental conditions during kernel maturation. In most dicotyledonous, and some monocotyledon seeds, the globulin types predominate in the grain. However, in the Triticeae (wheat, barley, rye, triticale) the major portion of seed proteins are not globulins but classes of protein characterized by regular repetitive domains with unique and fundamental functional features that determine wheat quality (Macritchie et al. 1990; Shewry and Halford 2002) and variation in these glutenin proteins either

quantitatively or qualitatively has major effects on end-product quality. The predictive power of defining the complement of glutenin subunit proteins for flour processing properties in breeding populations has been demonstrated for difficult traits such as extensibility (Eagles et al. 2002, 2004).

The physical properties of dough play a large role in determining its functionality. The extent to which a piece of dough can be stretched, and the force required to do so helps to determine the suitability of a variety for specific end-uses (Simmonds 1989; Eliasson and Larsson 1993). For example, leavened bread is best produced from dough that possesses strong and balanced rheological properties. In contrast, dough for biscuit production is generally less resistant to extension but is able to be extended a large distance before rupturing. This allows the dough to flatten and spread into a large, flat, evenly shaped biscuit (Simmonds 1989). Dough rheology is often measured using either an Extensograph (Brabender, Germany) or an Alveograph (Chopin). In both cases, slow sample throughput, the requirement for large sample sizes and the impact of extraneous error hamper genetic gain for improved rheological properties in wheat.

Strong dough will form a cohesive mass that has resistance to extension and is stable during mixing (Simmonds 1989). Such dough is able to hold the gas produced during fermentation within evenly distributed discrete cells and generates a loaf crumb with a much valued appearance and texture (Simmonds 1989). Such a crumb structure appears light in color, fine and silky in structure, both highly desirable quality attributes. Soft gluten will allow the gas cells to expand excessively during fermentation, causing their walls to collapse and the cells to coalesce together. The resulting bread has an undesirable, coarse, structure resulting from a very uneven texture (Simmonds 1989; Finney et al. 1987).

21.3.1 The Wheat Flour Proteins

Traditionally four different groups of proteins have been identified in wheat flour: (1) albumins (soluble in water and dilute buffers); (2) globulins (not soluble in water but soluble in saline solutions); (3) prolamins (which are soluble in 70–90% ethanol); (4) glutenins (which are soluble in dilute acid or alkali) (Beccari 1745; Osborne 1907). The glutenins derive from the gluten polymer and are usually subdivided into two distinct groups reflecting their solubility in 70% ethanol, namely glutenin and gliadins (Wrigley et al. 1996; Shewry and Casey 1999). Gliadins are usually single polypeptide chains (monomeric proteins) and the glutenins are multichained structures of polypeptides that are held together by disulfide bonds. The very high molecular weight of the polymeric structures formed by glutenins is responsible for their distinct contribution to the dough processing in contrast to gliadins which do not form large polymers (Fig. 21.4). Therefore, the classification of these proteins into monomeric and polymeric forms is a good indicator of dough functional properties (Payne and Lawrence 1983).

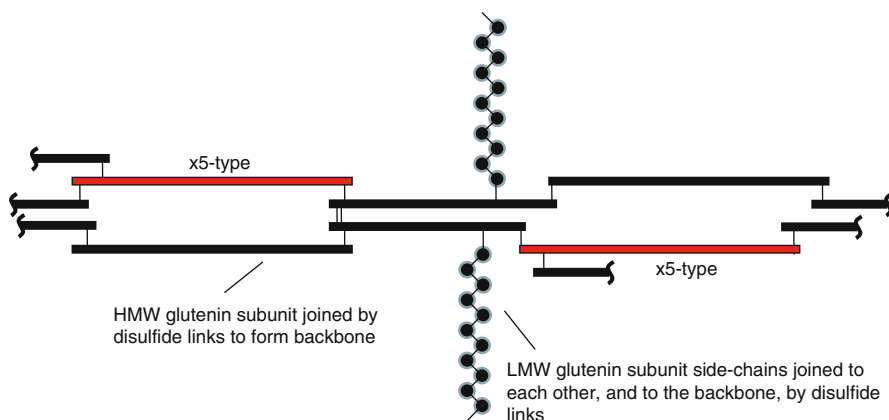


Fig. 21.4 Diagrammatic representation of the gluten polymer. The high molecular weight (HMW) glutenin subunit shown in red is particularly active in forming disulfide based polymer structures because it has an additional cysteine residue available for inter molecular linkages. The vertical entities in the figures are the low molecular weight (LWM) glutenin subunits. The diagram is based on Miles et al. (1991) and Wieser (2007) (See Color Insert)

The investigation of glutenin proteins in relation to dough properties have indicated two key variables, (1) the nature of the protein allele (reviewed in Gras et al. 2001) and (2) the level at which the respective allele is expressed (Butow et al. 2003). In some cases the control of the level of expression of seed storage protein has been well studied (reviewed in Appels et al. 2003; Ravel et al. 2006) but in general the interaction between promoters controlling gene expression and the environmental conditions during grain maturation have not been analyzed in any detail. The DNA structures of most of the glutenin genes have been determined and in some cases the genomic regions carrying the genes have been sequenced (Anderson et al. 2003; see also Section 21.3.2).

The prolamins forming the polymer are mainly the high- and low-molecular-weight glutenins, while the monomeric (polymer non-participating) prolamins are called α -, γ -, and ω -gliadins. In addition to the prolamins, other major wheat seed proteins include a wide assortment of proteins variously classified as globulins/albumins – including the starch granule associated grain-softness-protein and puroindolines, CM (chloroform-methanol soluble) proteins, amylase- and proteinase-inhibitors, thionins, and numerous other proteins of unknown function. Many of these additional seed proteins retain conserved motifs, such as cysteine numbers and placements, suggesting a common evolutionary source with the prolamins – including a number of proteins similar to gliadins but lacking the prominent repetitive domains characteristic of the wheat prolamins and tentatively named LMW-gliadins (Anderson et al. 2001; also referred to as foam stabilizing proteins, Clarke et al. 2001, 2003). All of these proteins have a common ancestry along with many dicot seed proteins and can be considered members of a large and diverse superfamily. Why the prolamins arose to prominence in the Triticeae is not

known but could be either an, as yet, unknown selective advantage or simply a random result of seed protein divergence.

Within the wheat seed prolamins, two evolutionary lines have resulted in the major prolamin classes, with both lines originating separately within the superfamily of related seed proteins. From one line the LMW-glutenins and gliadins emerged – sharing enough similarities to establish a common origin. The major distinguishing feature of the LMW-glutenins from the gliadins is the presence of two cysteine residues available for intermolecular disulfide bonds – the basis of incorporation into the gluten polymer. The gliadins are considered monomeric due to the general absence of free cysteines. In the cases of the α - and γ -gliadins, the even number of cysteines form intramolecular disulfide bonds, while the ω -gliadins generally lack cysteine residues altogether. Although not major participants in the gluten polymer, the monomeric gliadins contribute to dough physical/chemical properties by extensive hydrogen bonding with the gluten polymer (Macritchie et al. 1990; Shewry and Halford 2002).

Several gliadin genes have been reported to contain an uneven number of total cysteines, thus potentially leaving one available for intermolecular cross-linkages (Lindsay and Skerrett 2000); γ -gliadin (Sugiyama et al. 1986; Ferrante et al. 2006), and ω -gliadin (Altenbach and Kothari 2007) are examples. The effect of these gliadins on wheat quality parameters is unclear, but gliadins and LMW-glutenins with an uneven number of cysteines would be “chain-terminators” since they would lack a second free cysteine to continue the polymerization process (D’Ovidio and Masci 2004).

The second evolutionary line of wheat prolamins contains only the HMW-glutenins, and likely arose as a tandem duplication of a globulin gene, with one of the resulting genes evolving into a HMW-glutenin (Kong et al. 2003). A subsequent tandem duplication of the chromosome fragment containing an ancestral HMW-glutenin and globulin genes resulted in the x- and y-type HMW-glutenin genes.

The two HMW-glutenin genes have remained as two conserved genes, while the gliadins and LMW-glutenin genes have radiated into large multi-gene families of highly variable member copy numbers. Estimates of gliadin/LMW-glutenin gene family sizes have varied from 50 to 150 (Anderson et al. 1997) to 300 (Okita et al. 1985) for the α -gliadins, and 30–40 for the LMW-glutenins (Sabelli and Shewry 1991; Cassidy et al. 1998). Fewer estimates have been attempted for the γ - and ω -gliadin gene families, but estimates that are consistent with protein studies indicate 20–40 γ -gliadin genes and 8–15 ω -gliadin genes. Such estimates are confounded by a high percent of inactive pseudogenes (Anderson 1991; Anderson and Greene 1997; D’Ovidio and Masci 2004) – a characteristic of cereal prolamins.

The complication in estimating active gene numbers is particularly striking in the region of the B-genome ω -gliadins (Gao et al. 2007) where large numbers of gene fragments are found. In this region there are eight ω -gliadins sequences, only two of which appear to code for expressed proteins, with the other six either full-sized pseudo-genes or truncated fragments of ω -gliadin sequences.

In examining available LMW-glutenin sequences, Ikeda et al. (2002) proposed 12 distinct LMW-glutenin gene types. A careful examination of ESTs suggests restricting that further to perhaps only 9 active LMW-glutenin genes in hexaploid cultivars (O. Anderson, unpublished). In such estimates, identical gene sequences resulting from recent gene duplications are not accounted for. Together, such findings suggest reducing estimates of the active wheat gliadin/LMW-glutenin gene family sizes to one-quarter to one-half of estimates based solely on Southern hybridization experiments, to 20–70 for the α -gliadins, 8–15 for the γ -gliadins, 4–6 for the ω -gliadins, and 8–15 for the LMW-glutenins.

21.3.1.1 High Molecular Weight Glutenin Subunits (HMWGS)

The HMW glutenins are relatively minor components in terms of quantity, but they are major determinants of gluten elasticity through promoting the formation of larger glutenin polymers and thus are key factors for bread-making (Tatham et al. 1985). They are encoded by the *Glu-1* loci *Glu-A1*, *Glu-B1*, and *Glu-D1* that are located on the long arms of chromosome 1A, 1B and 1D, respectively. Each locus includes two genes linked together encoding two different types of HMW glutenin subunits, x - and y -type subunits (Payne et al. 1981, 1987; Shewry et al. 1992). The x -type subunits generally have a higher molecular weight than the y -type subunits. Payne and Lawrence (1983) summarized the number of alleles at the *Glu-1* loci: three allelic forms for *Glu-1A*, eleven alleles for *Glu-1B*, and six alleles for *Glu-1D*. Later, more alleles have been identified as reported by McIntosh et al. (2003); see also Liu et al. (2003); Sun et al. (2006). Although six genes exist for HMW glutenin subunits, most common wheat cultivars possess only three to five HMW glutenin subunits (one to three subunits in durum wheats) due to gene silencing. All hexaploid wheats contain at least the *Bx*, *Dx*, and *Dy* protein subunits in their endosperm, while most cultivars also contain a *By* subunit and a *Ax* subunit as well. The gene encoding the *Ay* subunit is usually silent.

The x and y subunits share a highly similar primary structure, which consists of a signal peptide (removed from the mature protein), a N-terminal domain, a central repetitive domain, and a C-terminal domain (Shewry and Halford 2002). Most of the reported x -type subunits possess four conserved cysteine residues (three in the N-terminal domain, one in the C-terminal domain), and the majority of the y -type subunits characterized thus far contain seven conserved cysteine residues (five in the N-terminal domain, one in the repetitive domain and one in the C-terminal domain, Shewry and Tatham 1997). These cysteine residues are involved in the formation of disulphide bonds within and between subunits and are thus important for the high order structure and the functionality of these proteins in shaping the elastic properties of the gluten complex in wheat dough (Shewry and Tatham 1997; see also Fig. 21.4). In both x - and y -type subunits, the repetitive domains are composed of short and repeated peptide (tripeptide, hexapeptide, nanopeptide) motifs, with the presence of the tripeptide motif being unique to the repetitive domain of x -type

subunits (Shewry and Tatham 1997). With unusually high content of glutamic acid, HMW glutenin subunits also have high contents of proline and glycine and low contents of lysine. The amino acid composition of HMW glutenin subunits reveals the hydrophilic nature of the central repetitive domain and the hydrophobic characteristics of the N- and C-terminal domains (Shewry et al. 1989). The proportion of the different amino acids is mainly defined by sequences of repeated polypeptide motifs. Variations on the consensus repeat sequences PGQGQQ and GYYPTSPQQ form >90% of the repetitive domain (Anderson and Green 1989; Shewry and Tatham 1997).

Variations of the number of cysteine residues of HMW glutenin subunits have been associated with different properties in bread-making. For example, the *Dx5* subunit has an extra cysteine residue located at the N-terminal part of its repetitive domain and this subunit has frequently been found to associate with improved processing quality in bread wheat varieties (Lafiandra et al. 1993; Gupta and MacRitchie 1994; see also Fig. 21.4). A novel variant of the *Ax2** subunit, *Ax2*B*, is found to contain an extra cysteine residue located in the middle of its repetitive domain (Juhász et al. 2003). This subunit exerts a positive effect on the gluten properties. In contrast, the *Bx14* and *Bx20* subunits have reduced numbers of cysteine residues in their N-terminal domains (Shewry et al. 2003).

Increased dough resistance can also derive from an increase in expression of the *Bx7* subunit (Butow et al. 2003; *Glu-B1* locus (*al* allele)) compared to alternative alleles at that locus (Eagles et al. 2004). Structurally, the gene conferring *Bx7* over-expression has an 18 bp insertion in the central repetitive domain relative to the normal *Bx7* gene, and this feature has been used to developed PCR markers for differentiating these two genotypes. Other allelic variant pairs had similar results: *Glu-B1* subunits 17 + 18 (strong) versus subunits 20x + 20y (weak), (Cornish et al. 2001).

21.3.1.2 Low Molecular Weight Glutenin Subunits (LMWGS)

Most of the LMW glutenin subunits are encoded by the complex *Glu-3* loci (*Glu-A3*, *Glu-B3* and *Glu-D3*) on the short arms of chromosomes 1A, 1B and 1D (Gupta and Shepherd 1990; Jackson et al. 1983), though other LMW-GS gene loci were also reported, such as *Glu-B2* and *Glu-B4* on chromosome 1B (Jackson et al. 1985; Liu and Shepherd 1995), *Glu-D4* on chromosome 1D and *Glu-D5* on chromosome 7D (Sreeramulu and Singh 1997). Six, eleven and five alleles defined by protein electrophoretic mobility have been confirmed at *Glu-A3*, *Glu-B3* and *Glu-D3* locus, respectively, in common wheat (Branlard et al. 2003; Gianibelli et al. 2001; Gupta and Shepherd 1990). Recently, some new alleles such as *Glu-B3m*, *Glu-B3n* (McIntosh et al. 2003) and *Glu-D3f* (International Maize and Wheat Improvement Center and Japan NARC, unpublished) have been designated.

Based on the first amino acid present in the N-terminal sequences of the proteins, seven main types of LMW glutenin subunits have been identified,

which are LMW-s starting with the sequence SHIPGL-, three LMW-m with N-terminal sequences of METSHIPGL-, METSRIPGL and METSCIPGL-respectively, and three types with N-terminal sequences resembling those of the α -, β - and γ -type gliadins (Kasarda et al. 1988; Tao and Kasarda 1989; Lew et al. 1992; Cloutier et al. 2001). LMW glutenin subunits were further classified into 12 groups by Ikeda et al. (2002) according to deduced amino acid sequences and in particular the number and position of cysteine residues available for inter-molecular disulphide bond formation (Shewry and Tatham 1997). More than 100 genes, partial genes and pseudogenes of the LMW-GS family have been cloned and sequenced from several common wheat cultivars (Pitts et al. 1988; Cloutier et al. 2001; Ikeda et al. 2002; Zhang et al. 2004). Hai et al. (2005) retrieved 69 known LMW-GS genes from GenBank and classified them into nine groups by the deduced amino acid sequence of the highly conserved N-terminal domain, and nine corresponding primer sets proved to be LMW-GS group-specific were established. Ikeda et al. (2006) also constructed 10 group-specific markers according to the published nucleotide sequences. Zhao et al. (2006, 2007) analysed a set of Australian wheat cultivars and identified 6 different gene sequences and 12 gene haplotypes at the *Glu-D3* locus.

21.3.2 Seed Storage Protein Gene Structure and Variation

In spite of the large amount of literature on the wheat prolamins, there is still an incomplete understanding of the chromosomal organization of these genes and loci, but what is known gives insight both to the prolamins genes and global characteristics of the wheat genome. The HMW-glutenins have the simplest organization since there are only two genes per genome, and this region is the best studied area of the wheat genome (Gu et al. 2006). The immediate gene complement around the HMW-glutenin *Glu-1* loci is shown in Fig. 21.5A. The two HMW-glutenin genes are separated by 51–185 kb in the genomes thus far examined, with the variation in distance due to completely different complements of nested transposons among the A, B, and D genomes (Gu et al. 2006). In only the D genome have the two HMW-glutenin genes at the *Glu-1D* loci from both *T. tauschii* and *T. aestivum*, been found on a single BAC (Anderson et al. 2003; Gao et al. 2007). For all cases of the *Glu-1A* and *Glu-1B* loci, the spacing was sufficient to require two or three BACs to confirm spanning the HMW-glutenin intergenic regions of the A and B genomes: tetraploid (Gu et al. 2004), hexaploid (Gu et al. 2006), and *T. monococcum* (Gu et al. unpublished). The pattern of genes is conserved in all genomes studied including diploid, tetraploid, and hexaploid wheats; i.e. the locus contains a receptor kinase, globulin, γ -type HMW-glutenin, remnants of a globulin, α -type HMW-glutenin, and a protein kinase (Fig. 21.5A).

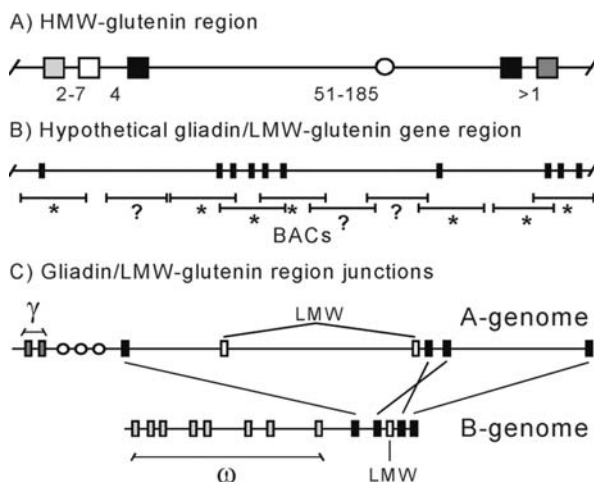


Fig. 21.5 Structure of wheat prolamins. The identity and relative spacing of examples of wheat prolamins is shown. (A) HMW-glutenin region. All wheat HMW-glutenin regions have the same gene order; i.e., a receptor kinase (light grey box), a globulin (white box), the γ -type HMW-glutenin (black box), fragments of a globulin (white circle), the α -type HMW-glutenin (black box), and a protein kinase (dark grey box). Variations in intervening sequences are shown in kilobases below gene pairs or between the two HMW-glutenins. (B) Generic gliadin/LMW-glutenin loci structure. A general model of the gliadin/LMW-glutenin loci structure is diagrammed to be small clusters of genes separated by longer sequences without prolamins. Possible BAC locations are shown under the loci to indicate how BACs isolated *via* prolamins probes may not span the entire loci space. Asterisks mark BACs containing prolamins. Question marks indicate BACs not containing prolamins and which span gaps in BAC contigs. (C) Comparing two BAC sequences containing common markers but different prolamins classes and gene orders. Common markers are black boxes. LMW-glutenin genes are white boxes. Dark grey boxes indicated γ -gliadin genes, and light grey boxes indicate ω -gliadin sequences. White circles are three genes related to the major prolamins but not classified

The gliadin/LMW-glutenin gene families are generally composed of larger numbers of genes – the exception being the *Glu-3A* locus, which contains 2–3 genes with only one gene likely to be functional (Wicker et al. 2003; Gao et al. 2007; Anderson unpublished). The other loci are structured with multiple genes spread over longer distances and with varying clustering of genes. Several studies have attempted, with limited success, to isolate BAC contigs over extended wheat genomic regions using specific known probes; i.e., LMW-glutenin/gliadins (Ozdemir and Cloutier 2005; Gao et al. 2007), and the *Ph1* region (Griffiths et al. 2006). In another study, four randomly chosen wheat BACs had gene densities of about 1 gene/75 kb (Devos et al. 2005). A similar average spacing was found when sequencing megabase-sized regions on the 3B chromosome (Choulet et al. 2008), and unpublished findings are that as many as one-third of wheat BACs are without genes. The many examples of relative clustering of small numbers of wheat genes into gene “islands”, the failure to

identify long BAC contigs using gene probes, and the known incidences of BAC+ sized regions absent of genes, together suggests the organization of the larger prolamin gene family regions is proposed as shown in Fig. 21.5B – islands of singlets and clusters of prolamin genes separated by stretches of repetitive DNA. As shown in Fig. 21.5B, probing with prolamin sequences will identify those BACs carrying prolamin genes and allow singleton/contig identification (BACs marked with asterisks in Fig. 21.5B) but would not identify intervening BACs without prolamin genes – resulting in a number of smaller contigs, but no single contig covering the entire locus. The larger of these loci could be very large. In the case of the α -gliadin gene family, previous results suggest the *Gli-3A* locus could contain as many as 100 genes (active genes plus pseudogenes). With an average gene spacing of 75 kb, this would make the *Gli-3A* locus 7–8 Mb in size in some cultivars.

Although there is limited data available on the DNA sequence structure adjoining prolamin loci, the available results suggest major differences among the wheat genomes. Similar to the study on the wheat α -gliadin loci by Gu et al. (2004), Gao et al. (2007) isolated LMW-glutenin, γ -gladin, and ω -gliadin containing BACs from tetraploid wheat. As with the α -gliadin study, contigs within each prolamin family showed small clustering of genes interspersed with longer DNA stretches generally precluding assembling complete contigs of these loci. However, two BACs, one each from A and B genomes, contained representatives of two different prolamin classes; i.e., LMW-glutenin and γ -gladin in one case, and LMW-glutenin and ω -gladin in the other. The initial assumption was that the junctions of these adjacent prolamin multigene families had been isolated with the two gliadin loci on opposite sides of the LMW-glutenin locus in the two genomes. Sequencing of the two BACs subsequently showed that the situation was in fact more complex than previously appreciated (Fig. 21.5C). Four markers are in common between the two sequences (black boxes and connecting lines in Fig. 21.5C), indicating common ancestry, but the order and identity of the other genes is different. In addition to the four common markers, the A-genome sequence contains the complete *Glu-A3* locus (two LMW-glutenin genes) plus two γ -gladin genes separated from the LMW-glutenin genes by three genes related to prolamins but not classifiable into the standard wheat prolamin classes. In the B-genome, the sequence includes a single LMW-glutenin bracketed by the common markers, with eight genes and gene fragments for ω -gliadins immediately adjacent – but no prolamin-related genes as in the A-genome. Figure 21.5C implies the contradiction that the two different gliadin loci are situated in the same orientation and adjacent with respect to the LMW-glutenin loci. The explanation may include features such as interspersion of the prolamin genes, local inversions, and other as yet unrecognized differences in organization.

21.3.2.1 Assaying Variation in Seed Storage Proteins

Glutenin alleles can be identified by extracting the glutenin proteins from seed and separating them on SDS PAGE gels (Singh et al. 1991). Although capable

of detecting a wide range of alleles at the six glutenin loci, this method is slow and expensive, limiting its use to advanced breeding material. MALDI-TOF based analysis of seed storage proteins is a technology that is capable of high resolution as well as high throughput (Zolla et al. 2002; Chen et al. 2007). Alternatively, DNA based molecular markers specific for particular glutenin alleles, have been developed to aid in selection for improved dough rheology (D'Ovidio et al. 1994; Devos et al. 1995; Ahmad 2000; Juhász et al. 2003; Radovanovic and Cloutier 2003; Zhang et al. 2004; Gale 2005). PCR-based assays to distinguish different HMW glutenins have been established. Sets of PCR primers can now amplify the complete coding region of the major HMW glutenin coding regions, including the *Ax*, *Bx*, *By*, *Dx* and *Dy* coding regions (D'Ovidio et al. 1994, 1995, 1996). PCR-based assays to specifically detect the *Dx5* gene (D'Ovidio and Anderson 1994) and to distinguish *Dy12* from *Dy10* (Smith et al. 1994) have been developed. Since *Dx5-Dy10* (*Glu-D1d*) and *Dx2-Dy12* (*Glu-D1a*) are the two predominant alleles found at the *Glu-D1* locus, the *Dx5* assay (D'Ovidio and Anderson 1994) has been adapted to permit rapid discrimination between these alleles (Varghese et al. 1996). De Bustos et al. (2000) used a PCR procedure based on minor nucleotide sequence variation of the regions immediately flanking the gene sequences of various HMW glutenin alleles to selectively amplify the entire coding regions of *Ax2@*, *Ax1* or *Ax* Null, *Dx5* and *Dy10*. PCR-based markers specific for the *Dx5* and *Dy12* genes and a co-dominant marker for distinguishing between *Bx7* and *Bx17* have been developed (Ahmad 2000). A multiplex PCR was performed with these primers but since the size of the codominant Bx marker was well above 2 kb, while markers for *Dx5* and *Dy12* were below 700 bp, the multiplexing was only successful with markers for *Dx5* and *Dy10*. A set of PCR markers targeting all three HMWGS loci were developed by Ma et al. (2003) and these markers were successfully amplified together to systematically identify major alleles at the three loci by one PCR reaction. To date, a range of PCR markers are available to discriminate individual HMW-GS genes, including *Dx5*, *Ax1*, *Bx17*, *Bx7*, *Bx7*(OE), *By8*, *By8**, *By18*, *20x+20y* (*e* allele), (Ma et al. 2003; Butow et al. 2003; Lei et al. 2006). Recently, efforts have been made to establish the relationship between different protein mobility alleles and their corresponding allelic variants at DNA level. Based on the allelic variation of one LMW-GS gene at *Glu-A3* locus, sets of PCR markers were developed by Zhang et al. (2004) to identify *Glu-A3* protein mobility alleles. Zhao et al. (2006, 2007) identified 12 gene haplotypes and developed 7 STS-markers for *gluD3* alleles.

21.3.3 Flour Color

Flour color is the result of the yellowness of the endosperm and the absence of finely divided bran specks; the latter reflects the ease with which the aleurone and outer layers can be removed from the endosperm.

21.3.3.1 The Yellowness of Flour and Its End Products

The level of yellow pigment is usually recorded as b^* using a Minolta meter, results largely from variation of xanthophyll (mainly lutein) levels in the grain (Mares and Mrva 2001). Yellow flour is undesirable for bread manufacture, while flour for noodle production can vary from a creamy to yellow color depending on the style of noodle (Simmonds 1989). Yellow alkaline noodles (YAN) require a bright, clear yellow color that develops by the addition of alkali, while white salted noodles (WSN) require a bright white color (Simmonds 1989). For white bread and many other bakery products, the flour needs to be white.

Major genes controlling xanthophyll content and therefore yellowness of flour are situated on chromosomes 7AL, 7BL, 7DL wheat (Parker et al. 1998; Ma 1999; Francki et al. 2004). Smaller less significant associations with flour yellowness have also been detected on chromosomes 3AS, 3BS (Parker et al. 1998; Mares and Mrva 2001; Francki et al. 2004) as well as a number of other individual chromosomes. One of the genes coding for phytoene synthase has been colocated to the QTL on the group 7 chromosomes (Pozniak et al. 2007; He et al. 2007b) and is thus a good candidate gene for the control of lutein levels in the endosperm since it is directly involved in the terminal pathway that generates lutein. The intron-exon structure for the phytoene synthase on chromosome 7A of wheat was studied in detail by He et al. (2007b) and polymorphisms in this gene accounted for 20–28% of variation in yellow pigment across three environments. A mutation in the phytoene synthase gene on 7E (long arm) of tall wheatgrass has been shown to be genetically linked to yellow pigment content (Zhang and Dubcovsky 2008), confirming the importance of the gene in this phenotype. Zhang and Dubcovsky (2008) have also provided evidence for an unidentified gene near the phytoene synthase gene on 7AL that contributes to the control of yellow pigment accumulation in durum wheat. A second phytoene synthase gene was located to chromosomes 5A and 5B by Pozniak et al. (2007) although no QTL for color has yet been found on these chromosomes. Thirteen other genes are involved at various points in the biosynthetic pathway to form lutein and the comparative genomics approach to studying these genes is ongoing (Francki et al. 2004).

21.3.3.2 The Finely Divided Bran Specks in Flour

The presence of finely divided bran specks in flour is an issue in wheat quality when these specks become visible to the naked eye as a result of polyphenol oxidase (PPO) present in the bran. Bread made from dark flour, caused by the presence of bran in the flour, will generally have a small loaf volume, a coarser texture and darker crumb colour than that made from white flour (Simmonds 1989). The breakdown of lipids through lipoyxygenase activity can also affect colour and cause off-flavours (Hessler et al. 2002); the *Lpx-B1* locus on chromosome 4BS is a major source of lipoyxygenase activity.

The PPO enzyme is widely distributed in plant species (Flurkey 1989). The PPO enzyme catalyzes hydroxylation of monophenols to o-diphenols and oxidation of o-diphenols to o-quinones that polymerize non-enzymatically into dark color matter (Okot-Kotber et al. 2002). High PPOs in grain and flour is responsible for the undesirable time-dependent darkening of bread wheat (*Triticum aestivum* L.) based end-products such as noodles (Kruger et al. 1994a,b; Baik et al. 1995; Crosbie et al. 1996; Anderson and Morris 2003). PPO activity varies among wheat genotypes and is also affected by environment (Baik et al. 1995; Park et al. 1997; Ge et al. 2003). Cultivars with low PPO activity are desirable for the consumers and food manufacturers.

PPO genes have been cloned and sequenced in several plant species (Bucheli et al. 1996; Thipyapong et al. 1997). Recently, efforts have been made to clone PPO genes from bread wheat to understand the molecular mechanism underlying darkening of wheat based end-products (Demeke and Morris 2002; Jukanti et al. 2004). The sequence information of PPO genes has been used by Demeke and Morris (2002) to design oligonucleotide primers from the conserved copper binding regions of other plant PPO genes and thus obtained putative DNA sequence for the wheat PPO gene (GenBank accession number AF507945). Anderson (2004) obtained a PPO gene sequence from wheat cDNA library (GenBank accession number AY515506). In addition, sequence information of some other PPO genes was obtained from full-length sequencing of EST clones (Jukanti et al. 2004, GenBank accession number AY596266, AY596267, AY596268, AY596269, and AY596270).

Udall (1997) identified a QTL for PPO activity in a recombinant inbred line population derived from a cross between NY18 and CC, and found a RFLP marker Xcdo373 on wheat chromosome 2A closely linked to the QTL accounting for over 40% of the variation of PPO activity. Jimenez and Dubcovsky (1999) reported that genes located in the wheat chromosome homoeologous group 2 played an important role in PPO activity. Demeke et al. (2001) used three inbred line populations to study wheat PPO genes distribution, chromosome location, and number of loci involved in wheat PPO, and found polygenic inheritance in two populations (M6/Opata85, NY18/CC) and monogenic inheritance in the third population (ND2603/Butte86). They identified a QTL significantly associated with wheat PPO activity on chromosome 2D in the M6/Opata85 mapping population. Raman et al. (2007) also found a major QTL for PPO activity on chromosome 2AL in a DH population derived from Chara/WW2449 and suggested that the SSR markers Xgwm294 and WMC170 may be used for marker-assisted selection. Zhang et al. (2005) detected a QTL on chromosome 2AL closely linked to Xgwm312 and Xgwm294, explaining 38% of the phenotypic variance of grain PPO activity.

PCR-based markers for the PPO genes on 2A and 2D have been published (Sun et al. 2005; Raman et al. 2007; Massa et al. 2007; He et al. 2007a); the marker for the 2A gene is particularly diagnostic and amplifies a 685-bp and a 876-bp fragment in the cultivars with high and low PPO activity, respectively. The difference of 191 bp is located in the intron region of the PPO genes.

21.3.4 Flour Paste Viscosity

The high paste viscosity of flour is critical in determining the quality of Japanese white salted (Oda et al. 1980; Konik and Miskelly 1992) and Chinese noodles (Miskelly and Moss 1985). In breeding programs, both the visco-analyser and flour swelling volume tests have been used to determine the viscosity and therefore quality of wheat for noodle manufacture (Crosbie 1991; Panozzo and McCormick 1993). Biochemically, the ratio of amylose to amylopectin in starch is one of the major determinants of this viscosity (Sasaki et al. 2000). Yamamori et al. (2000) showed the important positive correlation between the quantity of granule bound starch synthase (GBSS) in wheat flour and flour amylose content. The location of a homoeologous gene series (*Wx-A1*, *-B1* and *-D1*) on chromosomes 7A, 4A (ancient translocation from 7B) and 7D (Chao et al. 1989), and the subsequent development of molecular markers to aid in the selection of the null alleles at each of these loci (McLauchlan et al. 2001), has provided an important tool for the improvement of noodle quality. Relatively few reports (Udall et al. 1999; Igrejas et al. 2002) exist of genetic associations with flour viscosity, other than the association with the *Wx* gene series (see also Section 21.4).

21.4 Grain Hardness and Carbohydrates in Wheat and Barley

The importance of carbohydrates has become evident as some clear genetic variation in starch and non-starch carbohydrates is being identified. Due to the strong relationship between the starch granules and protein matrix, hard textured varieties suffer greater starch damage during milling than soft textured varieties. Soft textured varieties are better suited to biscuit, cookie and some noodle manufacture while hard grained varieties are used for bread and some noodles (Simmonds 1989). While grain hardness is correlated with grain protein content (Giroux et al. 2000; Martin et al. 2001), genetic control of texture independent of protein content also exists.

The importance of starch in determining human health outcomes from the consumption of cereal based foods has been increasingly recognized and, in addition, cereals have been increasingly used as sources of energy in biofuels and/or animal feed (Topping 2007; Rahman et al. 2007). This appreciation of the significance of the carbohydrates in cereals has led to investigating the complexities of starch and non-starch polysaccharide structure through identifying the genes involved in their synthesis (Martinant et al. 1998; Shewry and Morell 2001; Morell and Myers 2005). The manipulation of cereal carbohydrates is now widely seen as one of the key areas for future innovation in grain quality.

21.4.1 Starch Content

Reductions in starch content associated with starch compositional change can lead to potentially useful alterations in other grain components (e.g. in the *Sex6*

mutant of barley; Bird et al. 2004a,b; Morell et al. 2003; Topping et al. 2003) but in commercial applications this means there is a tradeoff between yield penalty and added value. While there is some natural genetic diversity in starch content known in small grain cereals, the major research focus has been on increasing the flux to starch by manipulating the properties of the first committed step in starch biosynthesis, namely, the step carried out by the enzyme ADP-glucose pyrophosphorylase (see Meyer et al. 2007; Sakulsingharoj et al. 2004). Traits of importance for high starch applications include feed wheat, wheats for bioethanol production, wheats for starch/gluten separation.

21.4.2 Starch Composition

The advent of improved techniques for the isolation of genes and the relationship of those genes to expressed proteins and biochemical functions, have led to the cloning and isolation of a core set of genes in cereal quality (Morell and Myers 2005). Genome sequencing brought access to genome sequences in *Arabidopsis* and rice that allowed the remaining genes and candidate genes considered to be important in starch biosynthesis to be identified. Two further techniques have been important in allowing the knowledge of gene sequences and their expressed proteins to be harnessed. Firstly, the development of gene silencing techniques, such as antisense and RNAi, allow the role of key genes in hexaploid wheat to be directly tested (Regina et al. 2006; Fu et al. 2007). Secondly, techniques exist to generate and discover diversity in gene sequences, allowing the development of waxy wheat (Nakamura et al. 1995), wheats with deficiency in starch synthase IIa (Yamamori et al. 2000) and wheats lacking functional starch branching enzyme I, BEI, genes from each wheat genome (Regina et al. 2004). Techniques such as TILLING (Slade et al. 2005) provide the potential for identifying and combining mutations in each of the starch biosynthetic genes in each genome of wheat, and hence a comprehensive resource for identifying and designing, starches ideally suited to particular applications (Morell and Myers 2005).

Traits of importance for triple null wheat germplasm include waxy wheats and high amylose wheats and their commercial potential is being explored. Although triple null genotypes can be extreme in starch properties such as very low gelling temperatures and have significant yield penalties, they can provide valuable directions for the utilization of single and double null mutations. The latter can have important economic advantages driven by subtle variations in starch properties, and the added advantage of minimal impact on starch content or yield. The well known example is that of Udon noodles where mutations in GBSS result in high value segregations (Zhao et al. 1996; see also Section 21.3.4). Recently Konik-Rose et al. (2007) have demonstrated that subtle but important effects, on starch properties are associated with mutations in starch synthase IIa, SSIIa.

21.4.3 Non-Starch Polysaccharides

Cell-wall polysaccharides have long been known to be important in cereal quality. The term “pentosans” has been coined to describe the mix of arabinoxylans in cereals and the use of enzymes modulating the level and polymerisation state of pentosans is an important aspect of the bread improver market. The levels of pentosans are also important in feed applications for monogastric animals, particularly poultry. In comparison to starch, less is known of the genetic determinants of cell wall polysaccharides, mainly because of the difficulties in using biochemical approaches to identify the key enzymes, and because many of the genes involved are members of the very complex “Cellulose-Like” (Csl) gene family (Burton et al. 2007; see also Fig. 21.1). The genes included in this group include those involved in synthesis of a range of polysaccharides in addition to (1,3;1,4)- β -D-glucan and arabinoxylan. Although key traits include control of carbohydrate digestibility (e.g. glycemic index, metabolisable energy), manipulation of pentosans is also expected to impact water absorption and baking quality (Shewry and Morell 2001; Howitt et al. 2003).

21.4.4 Grain Hardness

The major locus involved in the control of grain texture (Symes 1965), *Ha*, is localized to the short arm of chromosome 5D (Mattern et al. 1973; Sourdille et al. 1996). Subsequently, two closely linked genes encoding puroindoline proteins (Gautier et al. 1994) were identified in the same region and have been considered to most likely encode the variation in grain hardness associated with the *Ha* locus (Giroux and Morris 1997). Giroux and Morris (1998) suggested that a mutation in the puroindoline-b gene (*Pinb-D1*), leading to an amino acid change, results in altered protein structure and consequently the strength with which the puroindoline protein binds with membrane polar lipids. This in turn was considered to alter the strength of the bond between the starch granules and protein matrix. In Giroux and Morris (1998) a null allele was also identified at the other puroindoline gene, *Pina-D1*. Consequently, it was argued that if a variety possessed either of the mutant alleles (*Pina-D1b* or *Pinb-D1b*) at these loci, the resultant grain was hard in texture. Although a number of other alleles have now been detected at these loci (Morris 2002), three genotypes predominate in released cultivars; “soft”, *Pina-D1a/Pinb-D1a*; “hard”, *Pina-D1a/Pinb-D1b* and “extra hard”, *Pina-D1b/Pinb-D1a* (Cane et al. 2004). The water absorption of these three genotypic classes is positively correlated with grain hardness. In the work of Cane et al. (2004), varieties with the “extra hard” genotype absorbed 3.5% more water than varieties with the “hard” genotype and 8.3% more than those with the “soft” genotype. However the distinction between the water absorption of the “extra hard” and “hard” classes

was not observed by Martin et al. (2001). Both of these studies showed a drop in milling yield associated with the “extra hard” class. Beyond this major gene for grain texture, numerous QTL associated with grain hardness have been reported.

Only the 5D genome of hexaploid wheat has the *Ha* locus and the homologous loci on chromosomes 5A and 5B are absent. Consistent with this situation in hexaploid wheat, the *Ha* locus is also missing from the tetraploid progenitor (AABB), although present in the diploid progenitors – a major deletion event is therefore assumed to have occurred after the polyploidization event that generated the AABB tetraploid wheat. The *Ha* locus is defined by three genes, grain softness protein (*Gsp*), puroindoline a (*Pina*) and puroindoline b (*Pinb*) and extensive sequence analyses on the region were carried out by Chantret et al. (2004, 2005). Based on genomic DNA sequences identifiable in tetraploid wheat, the 5' boundary of the *Ha* locus was defined by the *Gsp* gene that is also present in the A, B genomes of tetraploid wheat. The 3' boundary was defined by a gene cluster (called *Gene7* and *Gene8*) also present in A, B genomes of tetraploid wheat. The *Ha* locus can therefore be defined by a ca 55 kb segment of genomic DNA containing the *Pina*, *Pinb*, two degenerate copies of *Pinb*, *Gene 3* (present only in the D genomes) and *Gene 5*. *Gene 3* and *Gene 5* were not annotated by Chantret et al. (2005). The study by Chantret et al. (2005) indicated major differences between the D genome progenitor locus and the D genome locus in hexaploid wheat and these included the deletion of about 38 kb of DNA sequence in the hexaploid locus relative to the diploid locus.

The analysis of the barley *Ha* locus (Caldwell et al. 2004) has indicated that some gene clusters such as the gene cluster 2, GC2 (annotated as *VAMP*, *GlcNAc*, *Gsp*) are conserved between wheat and barley. Rearrangements have however occurred in the locus and the puroindoline equivalents in barley (hordoinolines) are upstream from GC2 instead of downstream as in wheat and they are also in the opposite orientation. The interpretation of grain hardness in barley has not focused on milling attributes as has been the case for wheat but has instead focused on resistance of the grain to invasion by pests and diseases (Caldwell et al. 2004).

21.5 Traits that Are Not Analysed at the Genomic Level to Date

21.5.1 Milling Yield

Milling yield is the amount of flour that can be obtained from a given weight of grain. It reflects the amount of endosperm and the ease with which it can be separated from the germ and bran. The separation of endosperm as flour is considered to relate to the degree of cross-linking of cell wall components (arabinoxylans) and hence the intimacy of the bran-endosperm

interconnection. Milling yield is thus a measure of the actual amount of endosperm available, which is a very important component of the grain for commercial purposes (Simmonds 1989; Finney et al. 1987). Flour is extracted primarily from the endosperm of a wheat grain, and is composed of starch granules encased in a protein matrix. Surrounding the endosperm, the bran, along with the embryo (germ) forms the non-flour fraction of the grain. The quantity of flour able to be extracted from a grain is consequently a function of the bran to endosperm ratio, which is dictated by grain size and morphology, as well as the ease with which the endosperm is released from the bran (Simmonds 1989). Grain size and morphology characters are under both environmental and genetic control. A number of reports exist of genetic associations with milling yield and many have not been experimentally related to grain size and shape.

21.5.2 Water Absorption

The quantity of water absorbed by flour during dough formation varies between varieties. For bread products relatively high water absorption is required, whereas for biscuits and noodle production lower water absorption is desirable (Simmonds 1989). The water absorption of flour is heavily influenced by grain texture, grain protein content and the level of non-starch polysaccharides. Although the level of starch damage can be altered by the conditions used during milling, this attribute is related to grain texture and therefore under significant genetic control (as discussed in Section 21.4). Although non-starch polysaccharides have been shown to influence the water absorption properties of wheat flour (Shogren et al. 1987; Shewry and Morell 2001; Howitt et al. 2003), this has not yet been clearly demonstrated through genetic association studies.

21.5.3 Grain Protein Content

As protein is responsible for much of the functionality of flour and the quantity of protein within each grain forms a key quality criteria. The quantity of protein in a wheat kernel, expressed as a proportion of kernel weight, is heavily influenced by both the nitrogen and carbon supply to the developing grain. Consequently, a strong inverse relationship between grain yield and protein concentration exists (Cooper et al. 2001; Fabrizius et al. 1997). It would therefore be expected that many of the genes responsible for the grain yield and grain weight (Röder et al. 2007) of wheat would also influence grain protein content. However, genes that increase grain protein without reducing grain yield, would be of more interest to wheat breeders attempting to improve both grain yield and protein concentration simultaneously.

Protein content relates to the amount of gluten in a sample and is very important measure of rheological properties (Wall 1979; Payne et al. 1984). Different wheat products have their own specific protein content requirements. For examples, biscuit manufacture requires low protein content (<11%), while pasta manufacture requires a minimum of 12% protein content. Manufacturing YAN noodles needs at least 11% protein content, while WSN noodles requires 9–9.5% protein content (Simmonds 1989).

A gene influencing protein content, independent of grain yield, was identified in *Triticum turgidum* on chromosome 6BS (Joppa et al. 1997). This gene (*Gpc-B1*) was transferred to bread wheat, resulting in the variety “GluPro” (Khan et al. 2000). Since then, the gene has successfully been incorporated into commercial varieties such as “Lillian” (DePauw et al. 2005), and “Somerset” (Fox et al. 2006). The detailed analysis of the *Gpc-B1* gene (designated *TaNAM-B1* by Uauy et al. 2006; Distelfeld et al. 2007) indicates its fundamental role is to act as a transcription factor controlling the timing of leaf senescence and the remobilization of zinc, iron and manganese. The increase in grain protein appears to be a pleiotropic effect associated with increased remobilization of protein from the senescent leaves (Distelfeld et al. 2007).

Numerous reports (Bresseghello et al. 2005; Prasad et al. 2003; Groos et al. 2004; Turner et al. 2004) have identified QTL associated with protein content, and that these are distributed across most of the genome. Unfortunately, many of these studies were performed without accounting for grain yield, so it is difficult to determine if selection for the high protein alleles within a breeding program would result in an increase *per se* in protein content without a corresponding drop in grain yield. However, it is interesting to note the large number of grain and flour protein content QTL that are coincident with grain yield and grain weight QTL identified in alternative populations. This tends to confirm the strong relationship between protein content and grain yield.

21.6 Impact of New Technologies

The impact of technologies that drive the high throughput analysis of protein and carbohydrate components of the grain and the genes that underpin variation in these components, will open up new requirements for integrating large and complex datasets. In wheat flour, for example, 1,300–1,500 polypeptides have been identified using 2D electrophoresis (Skylas et al. 2001; Gobaa et al. 2007) and these represent approximately 20–25% of the total genes expressed in wheat endosperm during development, based on sequence analyses of expressed sequence tags (ESTs) randomly cloned from mid-development endosperm tissue (Clarke et al. 2000). Since not all the genes expressed during development contribute to proteins in the mature kernel, it is evident that carrying out the appropriate proteomics analysis can identify a significant proportion of the genes contributing to protein components of wheat flour. The advantage of this

level of analysis, to complement the work ongoing using DNA probes, is that the effects of environment on protein accumulation during grain filling is also assessed as well as identifying the particular allele of the protein. Aspects of the allergic response of humans to wheat and barley products such as celiac disease have been reviewed (Kasarda 1994; van Heel and West 2006) and identifying key components responsible for these medical challenges is now becoming possible with the new analytical technologies.

Rapid advances in separation technologies such high performance liquid and gas chromatography linked to mass spectrometers can now define the metabolomics area of grain quality. Fast screens for vitamin A, iron and zinc in food products are possible in order to screen for genetic variants of grain that have higher levels of these compounds (HarvestPlus biofortification program; www.harvestplus.org). Similarly screening for low phytate levels is also important in the context of improving the nutritional attributes of grain because phytate interferes with the absorption of these and other minerals by the human gut.

The datasets derived from proteomics and metabolomics interface with the datasets from genome sequencing studies and microarray analyses of expressed genes in developing grain (Druka et al. 2006), germinating grain (Potokina et al. 2002, 2006) and plant responses to disease and abiotic stress (Walia et al. 2006), and web-based technologies form the basis for integrating these datasets. Assigning functions to specific genes is now also more efficient in cereals through the use of small RNAs or TILLING, to delete the activity of target genes through specifically removing messenger RNA or identifying mutations, respectively (Fu et al. 2007; Slade et al. 2005). A published example of a large-scale integration project is the Physiome project which is compiling all the information related to the human body (Hunter and Borg 2003) and this provides a good guide for dealing with specific areas such as grain quality evolving in-step with other changes in the food chain.

21.7 Conclusions

The quality attributes of cereal grains are valued in the context of a complex food chain that integrates outputs achievable by breeding, production, processing and consumer preferences. Each of these areas are constantly changing either in a controlled way (introduction of new technology) or in an uncontrolled way due to environmental variables and climate change, as well as changes in the market place resulting from advances in large-scale processing equipment and changes in consumer preferences. There is therefore a need for the quality features of wheat and barley grain to be tailored to keep step with these changes. The advances in the genomics of quality described in this chapter provide the basis for ensuring that the genetic approaches encompassing the complexities of the gene networks underpinning quality attributes can meet the challenges presented by the rapid changes occurring within the food chain.

References

- Ahmad, M. (2000) Molecular marker-assisted selection of HMW glutenin alleles related to wheat bread quality by PCR-generated DNA markers. *Theor. Appl. Genet.* 101, 892–896.
- Altenbach, S.B. and Kothari, K.M. (2007) Omega gliadin genes expressed in *Triticum aestivum* cv. Butte 86: effects of post-anthesis fertilizer on transcript accumulation during grain development. *J. Cereal Sci.* 46, 169–177.
- Anderson, J.V. and Morris, C.F. (2003) Purification and analysis of wheat grain polyphenol oxidase (PPO) protein. *Cereal Chem.* 80, 135–143.
- Anderson, O.D. (1991) Characterization of members of a pseudogene subfamily of the wheat α -gliadin storage protein genes. *Plant Mol. Biol.* 16, 335–337.
- Anderson, O.D. and Greene, F.C. (1989) The characterization and comparative analysis of high-molecular-weight glutenin genes from genomes A and B of a hexaploid bread wheat. *Theor. Appl. Genet.* 77, 689–700.
- Anderson, O.D. and Greene, F.C. (1997) The α -gliadin gene family. II. DNA and protein sequence variation, subfamily structure and origins of pseudogenes. *Theor. Appl. Genet.* 95, 59–65.
- Anderson, O.D., Hsia, C.C., Adalstein, A.E., Lew, E.J.-L. and Kasarda, D.D. (2001) Identification of several new classes of low-molecular weight wheat gliadin-related proteins and genes. *Theor. Appl. Genet.* 103, 307–315.
- Anderson, O.D., Litts, J.C. and Greene, F.C. (1997) The α -gliadin gene family. I. Characterization of ten new wheat α -gliadin gene clones, evidence for limited sequence conservation of flanking DNA, and Southern analysis of the gene family. *Theor. Appl. Genet.* 95, 50–58.
- Anderson, O.D., Rausch, C., Moullet, O. and Lagudah, E.S. (2003) The wheat D-genome HMW glutenin locus: BAC sequencing, gene distribution and retrotransposon clusters. *Funct. Integr. Genomics* 3, 56–68.
- Appleford, N.E.J., Evans, D.J., Lenton, J.R., Gaskin, P., Croker, S.J., Devos, K.M., Phillips, A.L. and Hedden, P. (2006) Function and transcript analysis of gibberellin-biosynthetic enzymes in wheat. *Planta* 223(3), 568–582.
- Appels, R., Francki, M. and Chibbar, R. (2003) Advances in cereal functional genomics. *Funct. Integr. Genomics* 3, 1–24.
- Baik, B.K., Czuchajowska, Z. and Pomeranz, Y. (1995) Discoloration of dough for oriental noodles. *Cereal Chem.* 72, 198–205.
- Bailey, P.C., McKibbin, R.S., Lenton, J.R., Holdsworth, M.J., Flintham, J.E. and Gale, M. D. (1999) Genetic map locations for orthologous Vp1 genes in wheat and rice. *Theor. Appl. Genet.* 98(2), 281–284.
- Barr, A.R., Karakousis, A., Lance, R.C.M., Logue, S.J., Manning, S., Chalmers, K.J., Kretschmer, J.M., Boyd, W.J.R., Collins, H.M., Roumeliotis, S., Coventry, S.J., Moody, D.B., Read, B.J., Poulsen, D., Li, C.D., Platz, G.J., Inkerman, P.A., Panozzo, J.F., Cullis, B.R., Smith, A.B., Lim, P. and Langridge, P. (2003) Mapping and QTL analysis of the barley population Chebec x Harrington. *Aust. J. Agric. Res.* 54, 1125–1130.
- Beccari, J.B. (1745) *De Frumento. De Bononiensi Scientarium et Artium. Instituto atque Academia Commentarii.* Bologna 2, 122–127.
- Bezant, H., Laurie, D.A., Pratchett, N., ChoJ.ecki, J. and Kearsey, M.J. (1997) Mapping of QTL controlling NIR predicted hot water extract and grain nitrogen content in a spring barley cross using marker-regression. *Plant Breeding* 116, 141–145.
- Bird, A.R., Flory, C., Davies, D.A., Usher, S. and Topping, D.L. (2004a) A novel barley cultivar (Himalaya 292) with a specific gene mutation in starch synthase IIa raises large bowel starch and short-chain fatty acids in rats. *J. Nutr.* (April) 134, 831–835.
- Bird, A.R., Jackson, M., King, R.A., Davies, D.A., Usher, S. and Topping, D.L. (2004b) A novel high-amylose barley cultivar (*Hordeum vulgare* var. Himalaya 292) lowers plasma cholesterol and alters indices of large-bowel fermentation in pigs. *Br. J. Nutr.* (October), 92, 607–615.

- Bonnardeaux, Y., Li, C., Lance, R., Zhang, X.Q., Sivasithamparam, K. and Appels, R. (2008) Seed dormancy in barley: identifying superior genotypes through incorporating epistatic interactions. *Aust J. Agr. Res.* 59, 517–526.
- Branlard, G., Dardevet, M., Amiour, N. and Igrejas, G. (2003) Allelic diversity of HMW and LMW glutenin subunits and omega-gliadins in French bread wheat (*Triticum aestivum* L.). *Genet. Reso. Crop Evol.* 50, 669–679.
- Breseghello, F., Finney, P.L., Gaines, C., Andrews, L., Tanaka, J., Penner, G. and Sorrells, M.E. (2005) Genetic loci related to kernel quality differences between a soft and hard wheat cultivar. *Crop Sci.* 45, 1685–1695.
- Bucheli, C.S., Dry, I.B. and Robinson, S.M. (1996) Isolation of a full-length cDNA encoding polyphenol oxidase from sugarcane, a C4 grass. *Plant Mol. Biol.* 3, 1233–1238.
- Burton, R.A., Wilson, S.M., Hrmova, M., Harvey, A.J., Shirley, N.J., Medhurst, A., Stone, B.A., Newbigin, E.J., Bacic, A. and Fincher, G.B. (2006) Cellulose synthase-like CslF genes mediate the synthesis of cell wall (1,3;1,4)-beta-D-glucans. *Science* 311, 1940–1942.
- Butow, B.J., Ma, W., Gale, K.R., Cornish, G.B., Rampling, L., Larroque, O.R., Morell, M.K. and Bekes, F. (2003) Molecular discrimination of Bx7 alleles demonstrates that a highly expressed high-molecular weight glutenin allele has a major impact on wheat flour strength. *Theor. Appl. Genet.* 107, 1524–1532.
- Caldwell, K.S., Langridge, P. and Powell, W. (2004) Comparative sequence analysis of the region harboring the hardness locus in barley and its colinear region in rice. *Plant Physiol.* 136, 3177–3190.
- Cane, K., Spackman, M. and Eagles, H.A. (2004) Puroindoline genes and their effects on grain quality traits in southern Australian wheat cultivars. *Aust. J. Agric. Res.* 55, 89–95.
- Cassidy, B.G., Dvorak, J. and Anderson, O.D. (1998) The wheat low-molecular-weight glutenin genes: characterization of six new genes and progress in understanding gene family structure. *Theor. Appl. Genet.* 96, 743–750.
- Chantret, N., Cenci, A., Sabot, F., Anderson, O. and Dubcovsky, J. (2004) Sequencing of the *Triticum monococcum* Hardness locus reveals good microcolinearity with rice. *Mol. Genet. Gen.* 271, 377–386.
- Chantret, N., Salse, J., Sabot, F., Rahman, S., Bellec, A., Laubin, B., Dubois, I., Dossat, C., Sourdille, P., Joudrier, P., Gautier, M.F., Cattolico, L., Beckert, M., Aubourg, S., Weissenbach, J., Caboche, M., Bernard, M., Leroy, P. and Chalhoub, B. (2005) Molecular basis of evolutionary events that shaped the hardness locus in diploid and polyploid wheat species (*Triticum* and *Aegilops*). *Plant Cell* 17, 1033–1045.
- Chao, S., Sharp, P.J., Worland, A.J., Warham, E.J., Koebner, R.M.D. and Gale, M.D. (1989) RFLP-based genetic maps of wheat homoeologous group 7 chromosomes. *Theor. Appl. Genet.* 78, 495–504.
- Chen, J., Lan, P., Tarr, A., Yan, Y.M., Francki, M., Appels, R. and Ma, W. (2007) MALDI-TOF based wheat gliadin protein peaks are useful molecular markers for wheat genetic study. *Rapid Comm. Mass Spectrom.* 21, 2913–2917.
- Choulet, F., Wicker, T., Paux, E., Sainetnac, C., Sourdille, P., Keller, B., Appels, R. and Feuillet, C. (2008) Plant and Animal Genome XVI Conference w278, www.intl-pag.org
- Clarke, B.C., Hobbs, M., Skylas, D.J. and Appels, R. (2000) Genes active in developing endosperm. *Funct. Integr. Genomics* 1, 44–55.
- Clarke, B.C., Larroque, O.R., Bekes, F. and Appels, D. (2001) The frequent classes of expressed genes in wheat endosperm tissue as possible sources of genetic markers. *Aust. J. Agric. Res.* 52, 1181–1193.
- Clarke, B.C., Phongkham, T., Gianibelli, M.C., Beasley, H. and Békés, F. (2003) The characterisation and mapping of a family of LMW-gliadin genes: effects on dough properties and bread volume. *Theor. Appl. Genet.* 106, 629–635.
- Cloutier, S., Rampitsch, C., Penner, G.A. and Lukow, O.M. (2001) Cloning and expression of a LMW-i glutenin gene. *J. Cereal Sci.* 33, 143–154.

- Collins, H.M., Panozzo, J.F., Logue, S.J., Jefferies, S.P. and Barr, A.R. (2003) Mapping and validation of chromosome regions associated with high malt extract in barley (*Hordeum vulgare* L.). *Aust. J. Agric. Res.* 54, 1223–1240.
- Cooper, M., Woodruff, D.R., Phillips, I.G., Basford, K.E. and Gilmour, A.R. (2001) Genotype-by-management interactions for grain yield and grain protein concentration of wheat. *Field Crops Res.* 69, 47–67.
- Cornish, G.B., Skylas, D.J., Siriamornpun, S., Bekes, F., Larroque, O.R., Wrigley, C.W. and Wootton, M. (2001) Grain proteins as markers of genetics traits in wheat. *Aust. J. Agric. Res.* 52, 1161–1171.
- Crosbie, G.B. (1991) The relationship between starch swelling properties, paste viscosity and boiled noodle quality in wheat flours. *J. Cereal Sci.* 13, 145–150.
- Crosbie, G.B., Solah, V.A., Chiu, P. and Lambe, W.J. (1996) Selection for improved color stability in noodles. In: C.W. Wrigley (Ed.), *Proceedings of the Australian Cereal Chemistry Conference 46th*, Sydney, Royal Australian Chemist Institute, North Melbourne, VIC, Australia, pp. 120–122.
- D'Ovidio, R. and Anderson, O.D. (1994) PCR analysis to distinguish between alleles of a member of a multigene family correlated with wheat bread-making quality. *Theor. Appl. Genet.* 88, 759–763.
- D'Ovidio, R., Lafiandra, D. and Porceddu, E. (1996) Identification and molecular characterization of a large insertion within the repetitive domain of a high-molecular-weight glutenin subunit gene from hexaploid wheat. *Theor. Appl. Genet.* 93, 1048–1053.
- D'Ovidio, R. and Masci, S. (2004) The low-molecular-weight glutenin subunits of wheat gluten. *J. Cereal Sci.* 39, 321–339.
- D'Ovidio, R., Masci, S. and Porceddu, E. (1995) Development of a set of oligonucleotide primers specific for genes at the *Glu-1* complex loci of wheat. *Theor. Appl. Genet.* 91, 189–194.
- D'Ovidio, R., Porceddu, E. and Lafiandra, D. (1994) PCR analysis of genes encoding allelic variants of high-molecular-weight glutenin subunits at the *Glu-D1* locus. *Theor. Appl. Genet.* 88, 175–180.
- De Bustos, A., Rubio, P. and Jouve, N. (2000) Molecular characterisation of the inactive allele of the gene *Glu-A1* and the development of a set of AS-PCR markers for HMW glutenins of wheat. *Theor. Appl. Genet.* 100, 1085–1094.
- DePauw, R.M., Townley-Smith, T.F., Humphreys, G., Knox, R.E., Clarke, F.R. and Clarke, J.M. (2005) Lillian hard red spring wheat. *Can. J. Plant Sci.* 85, 397–401.
- Demeke, T. and Morris, C.F. (2002) Molecular characterization of wheat polyphenol oxidase (PPO). *Theor. Appl. Genet.* 104, 813–818.
- Demeke, T., Morris, C.F., Campbell, K.G., King, G.E., Anderson, J.A. and Chang, H.-G. (2001) Wheat polyphenol oxidase distribution and genetic mapping in three inbred line populations. *Crop Sci.* 41, 1750–1757.
- Devos, K.M., Dubcovsky, J., Dvůrák, J., Chinoy, C.N. and Gale, M.D. (1995) Structural evolution of wheat chromosomes 4A, 5A, and 7B and its impact on recombination. *Theor. Appl. Genet.* 91, 282–288.
- Devos, K.M., Ma, J., Pontaroli, A.C., Pratt, L.H. and Bennetzen, J.L. (2005) Analysis and mapping of randomly chosen bacterial artificial chromosome clones from hexaploid bread wheat. *Proc. Natl. Acad. Sci. USA* 102, 19243–19248.
- Distelfeld, A., Cakmak, I., Peleg, Z., Ozturk, L., Yazici, A.M., Budak, H., Saranga, Y. and Fahima, T. (2007) Multiple QTL effects of wheat *Gpc-B1* locus on grain protein and micronutrient concentrations. *Physiologia Plantarum* 129, 635–643.
- Druka, A., Muehlbauer, G., Druka, I., Caldo, R., Baumann, U., Rostoks, N., Schreiber, A., Wise, R., Close, T., Kleinhofs, A., Graner, A., Schulman, A., Langridge, P., Sato, K., Hayes, P., McNicol, J., Marshall, D. and Waugh, R. (2006) An atlas of gene expression from seed to seed through barley development. *Funct. Integr. Genomics* 6, 202–211.

- Eagles, H.A., Eastwood, R.F., Hollamby, G.J., Martin, E.M. and Cornish, G.B. (2004) Revision of the estimates of glutenin gene effects at the *Glu-B1* locus from southern Australian wheat breeding programs. *Aust. J. Agric. Res.* 55, 1093–1096.
- Eagles, H.A., Hollamby, G.J., Gororo, N.N. and Eastwood, R.F. (2002) Estimation and utilization of glutenin gene effects from the analysis of unbalanced data from wheat breeding programs. *Aust. J. Agric. Res.* 53, 367–377.
- Eglinton, J.K. (2003) Novel alleles from wild barley for breeding malting barley (*Hordeum vulgare* L.). PhD thesis, University of Adelaide.
- Eglinton, J.K., Langridge, P. and Evans, D.E. (1998) Thermostability variation in alleles of barley beta-amylase. *J. Cereal Sci.* 28, 301–309.
- Eliasson, A. and Larsson, K. (1993) *Cereals in Breadmaking*. Marcel Dekker, New York.
- Fabrizius, M.A., Cooper, M. and Basford, K.E. (1997) Genetic analysis of variation for grain yield and protein concentration in two wheat crosses. *Aust. J. Agric. Res.* 48, 605–614.
- Ferrante, P., Masci, S., D'Ovidio, R., Lafiandra, D., Volpi, C. and Mattei, B. (2006) A proteomic approach to verify *in vivo* expression of a novel γ -gliadin containing an extra cysteine residue. *Proteomics* 6, 1908–1914.
- Finkelstein, R., Reeves, W., Ariizumi, T. and Steber, C. (2008) Molecular aspects of seed dormancy. *Ann. Rev. Plant Biol.* 59, 387–415.
- Finney, K.F., Yamazaki, W.T., Youngs, V.L. and Rubenthaler G.L. (1987) Quality of hard, soft, and durum wheats. In: E.G. Heyne (Ed.), *Wheat and Wheat Improvement*, 2nd ed., Agronomy Monograph 13. ASA, CSSA, and SSSA, Madison, WI, pp. 677–748.
- Flurkey, W.H. (1989) Polypeptide composition and amino-terminal sequence of broad bean polyphenoloxidase. *Plant Physiol.* 91, 481–483.
- Fox, S.L., Townley-Smith, T.F., Humphreys, D.G., McCallum, B.D., Fetch, T.G., Gaudet, D.A., Gilbert, J.A., Menzies, J.G., Noll, J.S. and Howes, N.K. (2006) Somerset hard red spring wheat. *Can. J. Plant Sci.* 86, 163–167.
- Francki, M., Carter, M., Ryan, K., Hunter, A., Bellgard, M. and Appels, R. (2004) Comparative organization of wheat homoeologous group 3S and 7L using wheat-rice synteny and identification of potential markers for genes controlling xanthophyll content in wheat. *Funct. Integr. Genomics* 4, 118–130.
- Fu, D., Uauy, C., Blechl, A. and Dubcovsky, J. (2007) RNA interference for wheat functional gene analysis. *Transgenic Res.* 16, 689–701.
- Gale, K.R. (2005) Diagnostic DNA markers for quality traits in wheat. *J. Cereal Sci.* 41, 181–192.
- Gali, J.V., Brown, C. and Wegener, M. (1998) The value of barley protein in livestock feeding in Queensland. *Aust. Agribusiness J.* 6, ISSN 1442-6951.
- Gao, S., Gu, Y.Q., Wu, J., Coleman-Derr, D., Huo, N., Crossman, C., Jia, J., Ren, Z., Anderson, O.D. and Kong, X. (2007) Rapid evolution and complex structural organization in genomics regions harboring multiple prolamin genes in the polyploid wheat genome. *Plant Mol. Biol.* 65, 189–203.
- Gao, W., Clancy, J.A., Han, F., Prada, D., Kleinhoffs, A. and Ullrich, S.E. (2003) Molecular dissection of a dormancy QTL region near the chromosome 7 (5H) L telomere in barley. *Theor. Appl. Genet.* 107, 552–559.
- Gautier, M.F., Aleman, M.E., Guirao, A., Marion, D. and Joudrier, P. (1994) *Triticum aestivum* puroindolines, two basic cystine-rich seed proteins: cDNA sequence analysis and development gene expression. *Plant Mol. Biol.* 25, 43–57.
- Ge, X.X., He, Z.H., Yang, J. and Zhang, Q.J. (2003) Polyphenol oxidase activities of Chinese winter wheat cultivars and correlations with quality characteristics. *Acta Agronomica Sinica* 29, 481–485.
- Gianibelli, M.C., Larroque, O.R., Macritchie, F. and Wrigley, C.W. (2001) Biochemical, genetic, and molecular characterization of wheat glutenin and its component subunits. *Cereal. Chem.* 78, 635–646.

- Giroux, M.J. and Morris, C.F. (1997) A glycine to serine change in puroindoline b is associated with grain hardness and low levels of starch-surface friabilin. *Theor. Appl. Genet.* 95, 857–864.
- Giroux, M.J. and Morris, C.F. (1998) Wheat grain hardness results from highly conserved mutations in the friabilin components puroindoline a and b. *Proc. Natl. Acad. Sci. USA* 95, 6262–6266.
- Giroux, M.J., Talbert, L.E., Habernicht, D.K., Lanning, S.P., Hemphill, A. and Martin, J.M. (2000) Association of puroindoline sequence type and grain hardness in hard red spring wheat. *Crop Sci.* 40, 370–374.
- Gobaa, S., Bancel, E., Kleijer, G., Stamp, P. and Branlard, G. (2007) Effect of the 1BL.1RS translocation on the wheat endosperm as revealed by proteomic analysis. *Proteomics* 7, 4349–4357.
- Grando, S. (2002) Food barley gains long-overdue attention, <http://www.icarda.org/Publications/Caravan/caravan16/focus/food.htm>
- Gras, P.W., Anderssen, R.S., Keentok, M., Bekes, F. and Appels, R. (2001) Gluten protein functionality in wheat flour processing: a review. *Aust. J. Agric. Res.* 53, 1311–1323.
- Griffiths, S., Sharp, R., Foote, T.N., Bertin, I., Wanous, M., Reader, S., Colas, I. and Moore, G. (2006) Molecular characterization of *Ph1* as a major chromosome pairing locus in polyploid wheat. *Nature* 439, 749–752.
- Groos, C., Bervas, E. and Charmet, G. (2004) Genetic analysis of grain protein content, grain hardness and dough rheology in a hard x hard bread wheat progeny. *J. Cereal Sci.* 40, 93–100.
- Gu, Y.Q., Kong, X., Luo, M., You, F.M., Coleman-Derr, D., Dubcovsky, J. and Anderson, O.D. (2004) Genomic organization of the complex alpha-gliadin loci in wheat. *Theor. Appl. Genet.* 109, 648–657.
- Gu, Y.Q., Salse, J., Coleman-Derr, D., Dupin, A., Crossman, C., Lazo, G.R., Huo, N., Belcram, H., Ravel, C., Charmet, G., Charles, M., Anderson, O.D. and Chalhou, B. (2006) Types and rates of sequence evolution at the high-molecular-weight glutenin locus in hexaploid wheat and its ancestral genomes. *Genetics* 174, 1–12.
- Gupta, R.B. and MacRitchie, F. (1994) Allelic variation at the glutenin subunit and gliadin loci, Glu-1, Glu-3 and Gli-1, of common wheats. II. Biochemical basis of the allelic effects on dough properties. *J. Cereal Sci.* 19, 19–29.
- Gupta, R.B. and Shepherd, K.W. (1990) Two-step one-dimensional SDS-PAGE analysis of LMW subunits of glutenin. I. Variation and genetic control of the subunits in hexaploid wheats. *Theor. Appl. Genet.* 80, 65–74.
- Hai, L., Yu, M.W., Ze, H.Y., Bernard, B., Eviatar, N. and You, L.Z. (2005) Classification of wheat low-molecular-weight glutenin sub-unit genes and its chromosome assignment by developing LMW-GS group-specific primers. *Theor. Appl. Genet.* 111, 1251–1259.
- Hamilton, D.M. and Lewis, M.J. (1974) Factors affecting wort extract and attenuation. *MBAA Tech. Quart.* 11, 31–35.
- Han, F., Ullrich, S.E., Clancy, J.A., Jitkov, V., Kilian, A. and Romagosa, I. (1996) Verification of barley seed dormancy loci via linked molecular markers. *Theor. Appl. Genet.* 92, 87–91.
- Han, F., Ullrich, S.E., Clancy, J.A. and Romagosa, I. (1999) Inheritance and fine mapping of a major barley seed dormancy QTL. *Plant Sci.* 143, 113–118.
- Hattori, T., Vasil, V., Rosenkrans, L., Hannah, C., McCarty, D.R. and Vasil, I.K. (1992) The *Viviparous-1* gene and abscisic acid activate the *C1* regulatory gene for anthocyanin biosynthesis during seed maturation in maize. *Genes Devel.* 6, 609–618.
- Hayes, P.M., Ceronio, J., Witsenboer, H., Kuiper, M., Zabeau, M., Sato, K., Kleinof, A., Kudrna, D., Killian, A., Saghai-Maroo, M. and Hoffman, D. (1997) The Snorth American Barley Genome Mapping Project Characterizing and exploiting genetic diversity and quantitative traits in barley (*Hordeum vulgare*) using AFLP markers. *J. Quant Trait Loci*, <http://probenalusda.gov.8000/other-docs/J.q1997-02/>

- Hayes, P.M., Castro, A., Corey, A., Filichkin, T., Johnson, M., Rossi, C., Sandoval, S., Vales, I., Vivar, H.E. and Von Zitzewitz, J. (2001) In: H.E. Vivar and A. McNab (Eds.), *Breeding Barley in the New Millennium*. CIMMYT, Mexico, pp. 47–60.
- Hayes, P.M. and Iyamabo, O. (1994) The North American Barley Genome Mapping Project. Summary of QTL effects in the Steptoe x Morex population. *Barley Genet. Newsltr.* 23, 98–143.
- Hayes, P.M., Liu, B.H., Knapp, S.J., Chen, F., Jones, B., Blake, T., Franckowiak, J., Rasmusson, D., Sorrells, M., Ullrich, S.E., Wesenberg, D. and Kleinhofs, A. (1993) Quantitative trait locus effects and environmental interaction in a sample of North American barley germplasm. *Theor. Appl. Genet.* 87, 392–401.
- He, X.Y., He, Z.H., Zhang, L.P., Sun, D.J., Morris, C.F., Fuerst, E.P. and Xia, X.C. (2007a) Allelic variation of polyphenol oxidase (PPO) genes located on chromosomes 2A and 2D and development of functional markers for the PPO genes in common wheat. *Theor. Appl. Genet.* 115, 47–58.
- He, X.Y., Zhang, Y.L., He, Z.H., Wu, Y.P., Xiao, Y.G., Ma, C.X. and Xia, X.C. (2007b) Characterization of phytoene synthase 1 gene (*Psy1*) located on common wheat chromosome 7A and development of a functional marker. *Theor. Appl. Genet.* 116, 213–221.
- Hejgaard, J., Rasmussen, S.K., Brandt, A. and Svendsen, I. (1985) Sequence homology between barley endosperm protein Z and protease inhibitors of the $\alpha 1$ -antitrypsin family. *FEBS Lett.* 180, 89–94.
- Hessler, T.G., Thomson, M.J., Bensch, D., Nachit, M.M. and Sorrells, M.E. (2002) Association of a lipoxygenase locus, *Lpx-B1*, with variation in lipoxygenase activity in durum wheat seeds. *Crop Sci.* 42, 1695–1700.
- Hobo, T., Kowyama, Y. and Hattori, T. (1999) A bZIP factor, TRAB1, interacts with VP1 and mediates abscisic acid-induced transcription. *Proc. Natl. Acad. Sci. USA* 96, 15348–15353.
- Howitt, C.A., Tamás, L., Solomon, R.G., Gras, P.W., Morell, M.K., Békés, F. and Appels, R. (2003) Modifying flour to target functionality to product attributes. In: S. Cauvain (Ed.), *Bread Making: Improving Quality*. Woodhead Publishing, Cambridge, UK, pp. 220–252.
- Hunter, P.J. and Borg, T.K. (2003) Integration from proteins to organs: the Physiome project. *Nat. Rev.* 4, 237–243.
- Igrejas, B., Gaborit, T., Oury, F., Chiron, H., Marion, D. and Branlard, G. (2002) Genetic and environmental effects on puroindoline-a and puroindoline-b content and their relationship to technological properties in French bread wheats. *J. Cereal Sci.* 34, 37–47.
- Ikeda, T.M., Araki, E., Fujita, Y. and Yano, H. (2006) Characterization of low-molecular-weight glutenin subunit genes and their protein products in common wheats. *Theor. Appl. Genet.* 112, 327–334.
- Ikeda, T.M., Nagamine, T., Fukuoka, H. and Yano, H. (2002) Identification of new low-molecular-weight glutenin subunit genes in wheat. *Theor. Appl. Genet.* 104, 1432–1442.
- Jackson, E.A., Holt, L.M. and Payne, P.I. (1983) Characterisation of high-molecular-weight gliadin and low-molecular-weight glutenin subunits of wheat endosperm by two-dimensional electrophoresis and chromosomal localisation of their controlling genes. *Theor. Appl. Genet.* 66, 29–37.
- Jackson, E.A., Holt, L.M. and Payne, P.I. (1985) *Glu-B2*, a storage protein locus controlling the D group of LMW glutenin subunits in bread wheat. *Genet. Res.* 46, 11–17.
- Jimenez, M. and Dubcovsky, J. (1999) Chromosome location of genes affecting polyphenol oxidase activity in seeds of common and durum wheat. *Plant Breeding* 118, 395–398.
- Joppa, L.R., Du, C., Hart, G.E. and Hareland, G.A. (1997) Mapping gene(s) for grain protein in tetraploid wheat (*Triticum turgidum* L.) using a population of recombinant inbred chromosome lines. *Crop Sci.* 37, 1586–1589.
- Juhász, A., Gárdonyi, M., Tamás, L. and Bedő, Z. (2003) Characterisation of the promoter region of *Glu-1Bx7* gene from overexpressing lines of an old Hungarian wheat variety.

- Proceedings of the Xth International Wheat Genetics Symposium, 1–6 September 2003, Paestum, Italy, pp. 1348–1350.
- Jukanti, A.K., Bruckner, P.L. and Fischer, A.M. (2004) Evaluation of wheat polyphenol oxidase genes. *Cereal Chem.* 81, 481–485.
- Karakousis, A., Barr, A.R., Chalmers, K.J., Ablett, G.A., Holton, T.A., Henry, R.J., Lim, P. and Langridge, P. (2003a) Potential of SSR markers for plant breeding and variety identification in Australian barley germplasm. *Aust. J. Agric. Res.* 54, 1197–1210.
- Karakousis, A., Barr, A.R., Kretschmer, J.M., Manning, S., Logue, S.J., Roumeliotis, S., Collins, H.M., Chalmers, K.J., Li, C.D., Lance, R.C.M. and Langridge, P. (2003b) Mapping and QTL analysis of the barley population Galleon x Haruna Nijo. *Aust. J. Agric. Res.* 54, 1131–1135.
- Kasarda, D.D. (1994) Defining cereals toxicity in coeliac disease. In: C. Feighery and F. O'Farrelly (Eds.), *Gastrointestinal Immunology and Gluten-Sensitive Disease*. Oak Tree Press, Dublin, pp. 203–220.
- Kasarda, D.D., Tao, H.P., Evans, P.K., Adalsteins, A.E. and Yuen, S.W. (1988) Sequencing of protein from a single spot of a 2-D gel pattern: N-terminal sequence of a major wheat LMW-glutenin subunit. *J. Exp. Bot.* 39, 899–906.
- Khan, M.A., Hussain, I. and Baloch, M.S. (2000) Wheat yield potential current status and future strategies. *Pak. J. Bio. Sci.* 3, 82–86.
- Kleinhofs, A. and Han, F. (2002) Molecular mapping of the barley genome. In: G.A. Slafer, J. L. Molina-Cano, R. Savin, J.L. Araus and I. Romagosa (Eds.), *Barley Science: Recent Advances from Molecular Biology to Agronomy of Yield and Quality*. Food Products Press, New York, pp. 31–63.
- Komatsuda, T., Pourkheirandish, M., He, C., Azhaguvel, P., Kanamori, H., Perovic, D., Stein, N., Graner, G., Wicker, T., Tagiri, A., Lundqvist, U., Fujimura, T., Matsuoka, M., Matsumoto, T. and Yano, M. (2007) Six-rowed barley originated from a mutation in a homeodomain-leucine zipper I-class homeobox gene. *Proc. Natl. Acad. Sci. USA* 104, 1424–1429.
- Kong, X., Gu, Y.Q., You, F.M., Dubcovsky, J. and Anderson, O.D. (2003) Dynamics of the evolution of orthologous and paralogous portions of a complex locus regions in two genomes of allopolyploid wheat. *Plant Mol. Biol.* 54, 56–69.
- Konik, C.M. and Miskelly, D.M. (1992) Contribution of starch and non-starch parameters to the eating quality of Japanese white salted noodles. *J. Sci. Food Agric.* 58, 403–406.
- Konik-Rose, C., Thistleton, J., Chanvriat, H., Tan, I., Halley, P., Gidley, M., Kosar-Hashemi, B., Wang, H., Larroque, O., Ikea, J., Mcmaugh, S., Regina, A., Rahman, S., Morell, M. and Li, Z. (2007) Effects of starch synthase IIa gene dosage on grain, protein and starch in endosperm of wheat. *Theor. Appl. Genet.* 115, 1053–1065.
- Kruger, J.E., Anderson, M.H. and Dexter, J.E. (1994a) Effect of flour refinement on raw Cantonese noodle color and texture. *Cereal Chem.* 71, 177–182.
- Kruger, J.E., Hatcher, D.W. and DePauw, R. (1994b) A whole seed assay for polyphenol oxidase in Canadian prairie spring wheats and its usefulness as a measure of noodle darkening. *Cereal Chem.* 71, 324–326.
- Lafiandra, D., D'Ovidio, R., Porceddu, E., Margiotta, B. and Colaprico, G. (1993) New data supporting high Mr glutenin subunit 5 as determinant of qualitative differences in the pairs 5+10 vs 2+12. *J. Cereal Sci.* 18, 197–205.
- Lee, S.C., Cheng, H., King, K.E., Wang, W.F., He, Y.W., Hussain, A., Lo, J., Harberd, N.P. and Peng, J.R. (2002) Gibberellin regulates Arabidopsis seed germination via RGL2, a GAI/RGA-like gene whose expression is up-regulated following imbibition. *Genes Dev.* 16(5), 646–658.
- Lei, Z.S., Gale, K.R., He, Z.H., Gianibelli, M.C., Larroque, O., Xia, X.C., Butow, B.J. and Ma, W.J. (2006) Y-type gene specific markers for enhanced discrimination of high-molecular-weight glutenin alleles at the *Glu-B1* locus in hexaploid wheat. *J. Cereal Sci.* 43, 94–101.
- Lerouxel, O., Cavalier, D.M., Liepman, A.H. and Keegstra, K. (2006) Biosynthesis of plant cell wall polysaccharides – a complex process. *Curr. Opin.* 9, 621–630.

- Lew, E.J.L., Kuzmicky, D.D. and Kasarda, D.D. (1992) Characterization of low molecular weight glutenin subunits by reversed-phase high-performance liquid chromatography, sodium dodecyl sulphate-polyacrylamide gel electrophoresis, and N-terminal amino acid sequencing. *Cereal Chem.* 69, 508–515.
- Li, C., Ni, P., Francki, M., Hunter, A., Zhang, Y., Schibeci, D., Li, H., Tarr, A., Wang, J., Cakir, M., Yu, J., Bellgard, M., Lance, R. and Appels, R. (2004) Genes controlling seed dormancy and pre-harvest sprouting in a rice-wheat-barley comparison. *Funct. Integr. Genomics* 4, 84–93.
- Lindsay, M.P. and Skerriitt, J.H. (2000) The glutenin macropolymer of wheat flour doughs structure-function perspectives. *Trends Food Sci. Technol.* 10, 247–253.
- Liu, C.Y. and Shepherd, K.W. (1995) Inheritance of B subunits of glutenin and ω - and γ -gliadins in tetraploid wheats. *Theor. Appl. Genet.* 90, 1149–1157.
- Liu, Z., Yan, Z., Wan, Y., Liu, K., Zheng, Y. and Wang, D. (2003) Analysis of HMW glutenin subunits and their coding sequences in two *Aegilops* species. *Theor. Appl. Genet.* 106, 1368–1378.
- Ma, W., Zhang, W. and Gale, K.R. (2003) Multiplex-PCR typing of high molecular weight glutenin alleles in wheat. *Euphytica* 134, 51–60.
- Ma, W.Q. (1999) Current status and evaluation of crop fertilization in Shandong Province. Ph.D. Thesis (in Chinese), China Agricultural University, Beijing, p. 116.
- Ma, Y., Eglinton, J.K., Evans, D.E., Logue, S.J. and Langridge, P. (2000) Removal of the four C-terminal glycine-rich repeats enhances thermostability and substrate binding affinity of barley β -amylase. *Biochemistry* 39, 13350–13355.
- Ma, Y.F., Evans, D.E., Logue, S.J. and Langridge, P. (2001) Mutations of barley β -amylase that improve its thermostability and substrate-binding affinity. *Mol. Gen. Genet.* 266, 345–352.
- MacGregor, A.W., Bazin, S.L., Macri, L.J. and Babb, J.C. (1999) Modelling the contribution of alpha-amylase, beta-amylase and limit dextrinase to starch degradation during mashing. *J. Cereal Sci.* 29, 161–169.
- MacGregor, A.W., LaBerge, D.E. and Meredith, O.S. (1971) Changes in barley kernels during growth and maturation. *Cereal Chem.* 48, 255–269.
- Macritchie, F., du Cros, D.L. and Wrigley, C.W. (1990) Flour polypeptides related to wheat quality. *Adv. Cereal Sci. Technol.* 10, 79–145.
- Mantovani, M.S., Bellini, M.F., Angeli, J.P.F., Oliveira, R.J., Silva, A.F. and Ribeiro, L.R. (2008) β -Glucans in promoting health: prevention against mutation and cancer. *Mutat. Res.-Rev. Mutat.* 658, 154–161.
- Mares, D.J. and Mrva, K. (2001) Mapping quantitative trait loci associated with variation in grain dormancy in Australian wheat. *Aust. J. Agric. Res.* 52, 1257–1265.
- Marquez-Cedillo, L.A., Hayes, P.M., Jones, B.L., Kleinhofs, A., Legge, W.G., Rossnagel, B. G., Sato, K., Ullrich, S.E. and Wesenberg, D.M. (2000) QTL analysis of malting quality in barley based on the doubled-haploid progeny of two elite North American varieties representing different germplasm groups. *Theor. Appl. Genet.* 101, 173–184.
- Martin, J.M., Froberg, R.C., Morris, C.F., Talbert, L.E. and Giroux, M.J. (2001) Milling and bread baking traits associated with puroindoline sequence type in hard red spring wheat. *Crop Sci.* 41, 228–234.
- Martinant, J.P., Cadalen, T., Billot, A., Chartier, S., Leroy, P., Bernard, M., Saulnier, L. and Branlard, G. (1998) Genetic analysis of water-extractable arabinoxylans in bread wheat endosperm. *Theor. Appl. Genet.* 97, 1069–1075.
- Massa, A.N., Beecher, B. and Morris, C.F. (2007) Phenol oxidase (PPO) in wheat and wild relatives: molecular evidence for a multigene family. *Theor. Appl. Genet.* 114, 1239–1247.
- Mather, D.E., Tinker, N.A., La Berge, D.E., Edney, M., Jones, B.L., Rossnagel, B.G., Legge, W.G., Briggs, K.G., Irvine, R.B., Falk, D.E. and Kasha, K.J. (1997) Regions of the genome that affect grain and malt quality in a North American two-row barley cross. *Crop Sci.* 37, 544–554.

- Mattern, P.J., Morris, R., Schmidt, J.W. and Johnson, V.A. (1973) Location of genes for kernal properties in wheat variety 'Cheyenne' using chromosome substitution lines. Proceedings of the 4th International Wheat Genetics Symposium, Columbia, Missouri, USA, pp. 803–708.
- McIntosh, R.A., Yamazaki, Y., Devos, K.M., Dubcovsky, J., Rogers, W.J. and Appels, R. (2003) Catalogue of gene symbols for wheat (MacGene 2003) [CD-ROM]. In: N.E. Pogna et al. (Eds.), Proceedings of the 10th International Wheat Genetics Symposium, Vol. 4, 1–6 September 2003, Pasetum, Italy. Istituto Sperimentale per la Cerealicoltura, Rome, Italy.
- McKibbin, R.S., Wilkinson, M.D., Bailey, P.C., Flintham, J.E., Andrew, L.M., Lazzeri, P.A., Gale, M.D., Lenton, J.R. and Holdsworth, M.J. (2002) Transcripts of Vp-1 homeologues are misspliced in modern wheat and ancestral species. *Proc. Natl. Acad. Sci. USA* 99, 10203–10208.
- McLauchlan, A., Ogonnaya, F.C., Hollingsworth, B., Carter, M., Gale, K.R., Henry, R.J., Holten, T.A., Morell, M.K., Rampling, L.R., Sharp, P.J., Shariflou, M.R., Jones, M.G. K. and Appels, R. (2001) Development of PCR-based DNA markers for each homoeo-allele of granule-bound starch synthase and their application in wheat breeding programs. *Aust. J. Agric. Res.* 52, 1409–1416.
- Meyer, F.D., Talbert, L.E., Martin, J.M., Lanning, S.P., Greene, T.W. and Giroux, M.J. (2007) Field evaluation of transgenic wheat expressing a modified Adp-glucose pyrophosphorylase large subunit. *Crop Sci.* 47, 336–342.
- Miles, M.J., Carr, H.J., McMaster, T.C., I'Anson, K.J., Belton, P.S., Morris, V.J., Field, J.M., Shewry, P.R. and Tatham, A.S. (1991) Scanning tunneling microscopy of a wheat storage protein reveals details of an unusual super-secondary structure. *Proc. Natl. Acad. Sci. USA* 88, 68–71.
- Miskelly, D.M. and Moss, H.J. (1985) Flour quality requirements for Chinese noodle manufacture. *J. Cereal Sci.* 3, 379–387.
- Morell, M., Kosar-Hashemi, B., Samuel, M.S., Chandler, P., Rahman, A., Buleon, A., Batey, I. and Li, Z. (2003) Identification of the molecular basis of mutations at the barley *Sex6* locus and their novel starch phenotype. *Plant J.* 33, 1–13.
- Morell, M.K. and Myers, A.M. (2005) Rational design of cereal starches. *Curr. Opin. Plant Biol.* 8, 204–210.
- Morris, C.F. (2002) Puroindolines: the molecular genetic basis of wheat grain hardness. *Plant Mol Biol.* 48, 633–647.
- Nakamura, T., Yamamori, M., Hirano, H., Hidaka, S. and Nagamine, T. (1995) Production of waxy (amylose-free) wheats. *Mol. Gen. Genet.* 248, 253–259.
- Oberthur, L., Blake, T.K., Dyer, W.E. and Ullrich, S.E. (1995) Genetic analysis of seed dormancy in barley (*Hordeum vulgare* L.). *J. Quant Trait Loci*, <http://www.ncgr.org/jag/>
- Oda, M., Yasuda, Y., Okazaki, S., Yamauchi, Y. and Yokoyama, Y. (1980) A method of flour quality assesment for Japanese noodles. *Cereal Chem.* 57, 253–254.
- Okita, T.W., Cheesbrough, V. and Reeves, C.D. (1985) Evolution and heterogeneity of the α - β -type and γ -type gliadin DNA sequences. *J. Biol. Chem.* 260, 8203–8213.
- Okot-Kotber, M., Liavoga, A., Yong, K.J. and Bagorogoza, K. (2002) Activation of polyphenol oxidase in extracts of bran from several wheat (*Triticum aestivum*) cultivars using organic solvents, detergents, and chaotropes. *J. Agric. Food Chem.* 50, 2410–2417.
- Osborne, T.B. (1907) The Protein of the Wheat Kernel. Publication No. 84. Carnegie Institute, Washington, DC.
- Ozdemir, N. and Cloutier, S. (2005) Expression analysis and physical mapping of low-molecular-weight glutenin loci in hexaploid wheat (*Triticum aestivum* L.). *Genome* 48, 401–410.
- Oziel, A., Hayes, P.M., Chem, F.Q. and Jones, B. (1996) Application of quantitative trait locus mapping to the development of winter-habit malting barley. *Plant Breeding* 115, 43–51.
- Palotta, M.A., Asayama, S., Reinheimer, J.M., Davies, P.A., Barr, A.R., Jeffries, S.P., Chalmers, K.J., Lewis, J., Collins, H.M., Roumeliotis, S., Logue, S.J., Coventry, S.J., Lance, R.C.M., Karakousis, A., Lim, P., Verbyla, A.P. and Eckermann, P.J. (2003)

- Mapping and QTL analysis of the barley population Amagi Nijo x W12585. *Aust. J. Agric. Res.* 54, 1141–1144.
- Panozzo, J.F. and McCormick, K.M. (1993) The rapid viscoanalyzer as a method of testing for noodle quality in a wheat breeding programme. *J. Cereal Sci.* 17, 25–32.
- Park, W.J., Shelton, D.R., Peterson, C.J., Martin, T.J., Kachman, S.D. and Wehling, R.L. (1997) Variation in polyphenol oxidase activity and quality characteristics among hard white wheat and hard red winter wheat samples. *Cereal Chem.* 74, 7–11.
- Parker, G.D., Chalmers, K.J., Rathjen, A.J. and Langridge, P. (1998) Mapping loci associated with flour colour in wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* 97, 238–245.
- Payne, P.I., Corfield, K.G. and Blackman, J.A. (1981) Correlation between the inheritance of certain high-molecular-weight subunits of glutenin and bread-making quality in progenies of six crosses of bread wheat. *J. Sci. Food Agric.* 32, 51–60.
- Payne, P.I., Holt, L.M., Jackson, E.A. and Law, C.N. (1984) Wheat storage proteins: their genetics and their potential for manipulation by plant breeding. *Philos. Trans. R. Soc. Lond. B* 304, 359–371.
- Payne, P.I. and Lawrence, G.J. (1983) Catalogue of alleles for the complex gene loci, *Glu-A1*, *Glu-B1*, and *Glu-D1* which code for the high-molecular-weight subunits of glutenin in hexaploid wheat. *Cereal Res. Commun.* 11, 29–35.
- Payne, P.I., Nightingale, M.A., Krattiger, A.F. and Holt, L.M. (1987) The relationship between HMW glutenin subunit composition and the bread-making quality of British-grown wheat varieties. *J. Sci. Food Agric.* 40, 51–65.
- Pitts, E.G., Rafalski, J.A. and Hedgecoth, C. (1988) Nucleotide sequence and encoded amino-acid sequence of a genomic gene region for a low-molecular-weight glutenin. *Nucleic Acids Res.* 16, 11376.
- Potokina, E., Prasad, M., Malysheva, L., Roeder, M.S. and Graner, A. (2006) Expression genetics and haplotype analysis reveal cis regulation of serine carboxypeptidase I (Cxp1), a candidate gene for malting quality in barley (*Hordeum vulgare* L.). *Funct. Integr. Genomics* 6, 25–35.
- Potokina, E., Sreenivasulu, N., Altschmied, L., Mickalek, W. and Graner, A. (2002) Differential gene expression during seed germination in barley (*Hordeum vulgare* L.). *Funct. Integr. Genomics* 2, 28–39.
- Powell, W., Thomas, W.T.B., Baird, E., Lawrence, P., Booth, A., Harrower, B., McNicol, J.W. and Waugh, R. (1997) Analysis of quantitative traits in barley by the use of amplified fragment length polymorphisms. *Heredity* 79, 48–59.
- Pozniak, C.J., Knox, R.E., Clarke, F.R. and Clarke, J.M. (2007) Identification of QTL and association of a phytoene synthase gene with endosperm colour in durum wheat. *Theor. Appl. Genet.* 114, 525–537.
- Prada, D., Ullrich, S.E., Molina-Cano, J.L., Cistue, L., Clancy, J.A. and Romagosa, I. (2004) Genetic control of dormancy in a Triumph/Morex cross in barley. *Theor. Appl. Genet.* 109, 62–70.
- Prasad, M., Kumar, N., Kulwal, P.L., Röder, M., Balyan, H.S., Dhaliwal, H.S. and Gupta, P. K. (2003) QTL analysis for grain protein content using SSR markers and validation of associated markers using NILs in bread wheat. *Theor. Appl. Genet.* 106, 659–667.
- Radovanovic, N. and Cloutier, S. (2003) Gene-assisted selection for high molecular weight glutenin subunits in wheat doubled haploid breeding programs. *Mol. Breeding* 12, 51–59.
- Rahman, S., Bird, A., Regina, A., Li, Z., Ral, J.P., Mcmaugh, S., Topping, D. and Morell, M. (2007) Resistant starch in cereals: exploiting genetic engineering and genetic variation. *J. Cereal Sci.* 46, 251–260.
- Raman, R., Raman, H. and Martin, P. (2007) Functional gene markers for polyphenol oxidase locus in bread wheat (*Triticum aestivum* L.). *Mol. Breeding* 19, 315–328.
- Ravel, C., Nagy, I.J., Martre, P., Sourdille, P., Dardevet, M., Balfourier, F., Pont, C., Giancola, S., Praud, S. and Charmet, G. (2006) Single nucleotide polymorphism, genetic

- mapping and expression of genes coding for the DOF wheat prolamin-box binding factor. *Funct. Integr. Genomics* 6, 310–321.
- Regina, A., Bird, A., Topping, D., Bowden, S., Freeman, J., Barsby, T., Kosar-Hashemi, B., Li, Z., Rahman, S. and Morell, M.K. (2006) High amylose wheat generated by RNA-interference improves indices of large bowel health in rats. *Proc. Nat. Acad. Sci. USA* 103, 3546–3551.
- Regina, A., Kosar-Hashemi, B., Li, Z., Rampling, L., Cmiel, M., Gianibelli, M.C., Konik-Rose, C., Larroque, O., Rahman, S. and Morell, M.K. (2004) Multiple isoforms of starch branching enzyme-I in wheat: lack of the major SBE-I isoform does not alter starch phenotype. *Funct. Plant Biol.* 31, 591–601.
- Röder, M.S., Huang, X-Q. and Börner, A. (2007) Fine mapping of the region on wheat chromosome 7D controlling grain weight. *Funct. Integr. Genomics* 8, 79–86.
- Sabelli, P.A. and Shewry, P.R. (1991) Characterization and organization of gene families at the *Gli-1* loci of bread and durum wheats by restriction fragment analysis. *Theor. Appl. Genet.* 83, 209–216.
- Sakulsingharoj, C., Choi, S.B., Hwang, S.K., Edwards, G.E., Bork, J., Meyer, C.R., Preiss, J. and Okita, T.W. (2004) Engineering starch biosynthesis for increasing rice seed weight: the role of the cytoplasmic Adp-glucose pyrophosphorylase. *Plant Sci.* 167, 1323–1333.
- Sasaki, T., Yasui, T. and Matsuki, J. (2000) Effect of amylase content on gelatinization, retrogradation, and pasting properties of starches from waxy and nonwaxy wheat and their F1 seeds. *Cereal Chem.* 77, 58–63.
- Shewry, P.R. and Casey, R. (Eds.). (1999) *Seed Proteins*. Kluwer, Dordrecht, pp. 1–10.
- Shewry, P.R. and Halford, N.G. (2002) Cereal seed storage proteins: structures, properties and role in grain utilization. *J. Exp. Bot.* 53, 947–958.
- Shewry, P.R., Halford, N.G. and Lafiandra, D. (2003) The genetics of wheat gluten proteins. In: J.C. Hall, J.C. Dunlap and T. Friedman (Eds.), *Advances in Genetics*, Vol. 49. Academic Press, San Diego, pp. 111–184.
- Shewry, P.R., Halford, N.G. and Tatham, A.S. (1989) The high molecular weight subunits of wheat, barley and rye: genetics, molecular biology, chemistry and role of wheat gluten structure and functionality. In: B.J. Mifflin (Ed.), *Oxford Surveys of Plant Molecular and Cell Biology*, Vol. 6. Oxford University Press, Oxford, pp. 163–219.
- Shewry, P.R., Halford, N.G. and Tatham, A.S. (1992) High-molecular-weight subunits of wheat glutenin. *J. Cereal Sci.* 15, 105–120.
- Shewry, P.R. and Morell, M. (2001) Manipulating cereal endosperm structure, development and composition to improve end-use properties. In: P.R. Shewry, P.A. Lazzeri, and K.J. Edwards (Eds.), *Advances in Botanical Research*, Vol. 34. Academic Press, San Diego, San Francisco, New York, Boston, London, Sydney, Tokyo, pp. 165–236.
- Shewry, P.R. and Tatham, A.S. (1997) Disulphide bonds in wheat gluten proteins. *J. Cereal Sci.* 25, 135–146.
- Shogren, M.D., Hashimoto, S. and Pomeranz, Y. (1987) Cereal pentosans: their estimation and significance. II. Pentosans and breadmaking characteristics of hard red winter wheat flours. *Cereal Chem.* 64, 35–41.
- Simmonds, D.H. (1989) Inherent quality factors in wheat. *Wheat and Wheat Quality in Australia*. Australian Wheat Board, pp. 31–61.
- Singh, N.K., Shepherd, K.W. and Cornish, G.B. (1991) A simplified SDS-PAGE procedure for separating LMW subunits of glutenin. *J. Cereal Sci.* 14, 203–208.
- Skylas, D.J., Copeland, L., Rathmell, W.G. and Wrigley, C.W. (2001) The wheat-grain proteome as a basis for more efficient cultivar identification *Proteomics* 1, 1542–1546.
- Slade, A.J., Fuerstenberg, S.I., Loeffler, D., Steine, M.N. and Facciotti, D.A. (2005) Reverse genetic, nontransgenic approach to wheat crop improvement by TILLING. *Nat. Biotechnol.* 23, 75–81.
- Smith, R.L., Schweder, M.E., Barnett, R.D. (1994) Identification of glutenin alleles in wheat and *triticale* using PCR-generated DNA markers. *Crop Sci.* 34, 1373–1378.

- Sourdille, P., Perretant, M.R., Charmet, G., Leroy, P., Gautier, M.F., Joudrier, P., Nelson, J.C., Sorrells, M.E. and Bernard, M. (1996) Linkage between RFLP markers and genes affecting kernel hardness in wheat. *Theor. Appl. Genet.* 93, 580–586.
- Sreeramulu, G. and Singh, N.K. (1997) Genetic and biochemical characterization of novel low molecular weight glutenin subunits in wheat (*Triticum aestivum* L.). *Genome* 40, 41–48.
- Sugiyama, T., Rafalski, A. and Söll, D. (1986) The nucleotide sequence of A wheat gamma-gliadin genomic clone. *Plant Sci.* 44, 205–209.
- Sun, D.J., He, Z.H., Xia, X.C., Zhang, L.P., Morris, C.F., Appels, R., Ma, W.J. and Wang, H. (2005) A novel STS marker for polyphenol oxidase activity in bread wheat. *Mol. Breeding* 16, 209–218.
- Sun, X., Hu, S., Liu, X., Qian, W., Hao, S., Zhang, A. and Wang, D. (2006) Characterization of the HMW glutenin subunits from *Aegilops searsii* and identification of a novel variant HMW glutenin subunit. *Theor. Appl. Genet.* 113, 631–641.
- Susuki, M., Ketterling, M.G., Li, Q.-B. and McCarty, D.R. (2003) *Viviparous1* alters global gene expression patterns through regulation of abscisic acid signaling. *Plant Physiol.* 132, 1664–1677.
- Symes, K.J. (1965) The inheritance of grain hardness in wheat as measured by the particle size index. *Aust. J. Agric. Res.* 16, 113–123.
- Taketa, S., Amano, S., Tsujino, Y., Sato, T., Saisho, D., Kakeda, K., Nomura, M., Suzuki, T., Matsumoto, T., Sato, K., Kanamori, H., Kawasaki, S. and Takeda, K. (2008) Barley grain with adhering hulls is controlled by an ERF family transcription factor gene regulating a lipid biosynthesis pathway. *Proc. Natl. Acad. Sci. USA* 105, 4062–4067.
- Tao, H.P. and Kasarda, D.D. (1989) Two-dimensional gel mapping and N-terminal sequencing of LMW-glutenin subunits. *J. Exp. Bot.* 40, 1015–1020.
- Tatham, A.S., Mifflin, B.J. and Shewry, P.R. (1985) The beta-turn conformation in wheat gluten proteins: relationship to gluten elasticity. *Cereal Chem.* 62, 405–442.
- Thipyapong, P., Joel, D.M. and Steffens, J.C. (1997) Differential expression and turnover of the tomato polyphenol oxidase gene family during vegetative and reproductive development. *Plant Physiol.* 113, 707–718.
- Thomas, W.T.B., Powell, W., Swanston, J.S., Ellis, R.P., Chalmers, K.J., Barua, U.M., Jack, P., Lea, V., Forster, B.P., Waugh, R. and Smith, D.B. (1996) Quantitative trait loci for germination and malting quality characters in a spring barley cross. *Crop Sci.* 36, 265–273.
- Topping, D. (2007) Cereal complex carbohydrates and their contribution to human health. *J. Cereal Sci.* 46, 220–229.
- Topping, D.L., Morell, M.K., King, R.A., Li, Z., Bird, A.R. and Noakes, M. (2003) Resistant starch and health – Himalaya 292, a novel barley cultivar to deliver benefits to consumers. *Starke* 55, 539–545.
- Turner, A.S., Bradburne, R.P., Fish, L. and Snape, J.W. (2004) New quantitative trait loci influencing grain texture and protein content in wheat. *J. Cereal Sci.* 40, 51–60.
- Uauy, C., Distelfeld, A., Fahima, T., Blechl, A. and Dubcovsky, J. (2006) A NAC gene regulating senescence improves grain protein, zinc and iron content in wheat. *Science* 314, 1298–1301.
- Udall, J. (1997) Important alleles for noodle quality in winter wheat as identified by molecular markers. M.S. Thesis, University of Idaho, Moscow, ID.
- Udall, J.A., Souza, E., Anderson, J., Sorrells, M.E. and Zemetra, R.S. (1999) Quantitative trait loci for flour viscosity in winter wheat. *Crop Sci.* 39, 238–242.
- Ullrich, S.E., Han, F. and Jones, B.L. (1997) Genetic complexity of the malt extract trait in barley suggested by QTL analysis. *J. Am. Soc. Brew. Chem.* 55, 1–4.
- van Heel, D.A. and West, J. (2006) Recent advances in coeliac disease. *Gut* 55, 1037–1046.
- Varghese, J.P., Struss, D. and Kazman, M.E. (1996) Rapid screening of selected European winter wheat varieties and segregating populations for the *Glu-D1d* allele using PCR. *Plant Breeding* 115, 451–454.

- Walia, H., Wilson, C., Wahid, A., Condamine, P., Cui, X. and Close, T.J. (2006) Expression analysis of barley (*Hordeum vulgare* L.) during salinity stress. *Funct. Integr. Genomics* 6, 143–156.
- Wall, J.S. (1979) The role of wheat proteins in determining baking quality. In: D.L. Laidman and R.G. Wyn Jones (Eds.), *Recent Advances in the Biochemistry of Cereals*. Academic Press, New York, pp. 275–311.
- White, C.L., Staines, V.E. and Staines, M.H. (2007) A review of the nutritional value of lupins for dairy cows. *Aust. J. Agric. Res.* 58, 185–202.
- Wicker, T., Yahiaoui, N., Guyot, R., Schlagenhauf, E., Liu, Z.D., Dubcovsky, J. and Keller, B. (2003) Rapid genome divergence at orthologous low molecular weight glutenin loci of the A and A^m genomes of wheat. *Plant Cell* 15, 1186–1197.
- Wieser, H. (2007) Chemistry of gluten proteins. *Food Microbiol.* 24, 115–119.
- Wrigley, C.W., Bushuk, W. and Gupta, R. (1996) Nomenclature: establishing a common gluten language. In: C.W. Wrigley (Ed.), *Gluten 96*. RACI, Melbourne, Australia, pp. 403–407.
- Yamamori, M., Fujita, S., Hayakawa, K., Matsuki, J. and Yasui, T. (2000) Genetic elimination of a starch granule protein, SGP-1, of wheat generates an altered starch with apparent high amylose. *Theor. Appl. Genet.* 101, 21–29.
- Yoshigi, N., Okada, Y., Maeba, H., Sahara, H. and Tamaki, T. (1995) Construction of a plastid used for the expression of a seven-fold mutant barley β -amylase with thermostability in *Escherichia coli* and properties of the sevenfold mutant β -amylase. *J. Biochem.* 118, 562–567.
- Zhang, L.P., Ge, X.X., He, Z.H., Wang, D.S., Yan, J., Xia, X.C. and Sutherland, M.W. (2005) Mapping QTLs for polyphenol oxidase activity in a DH population from common wheat. *Acta Agronomica Sinica* 31, 7–10.
- Zhang, W. and Dubcovsky, J. (2008) Association between allelic variation at the *Phytoene synthase 1* gene and yellow pigment content in the wheat grain. *Theor. Appl. Genet.* 116, 635–645.
- Zhang, W., Gianibelli, M.C., Rampling, L. and Gale, K.R. (2004) Characterisation and marker development for low molecular weight glutenin genes from *Glu-A3* alleles of bread wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* 108, 1409–1419.
- Zhao, X.C., Batey, I.L., Sharp, P.J., Crosbie, G., Barclay, I., Wilson, R., Morell, M.K. and Appels, R. (1996) A single genetic locus associated with starch granule and noodle quality in wheat. *J. Cereal Sci.* 27, 7–13.
- Zhao, X.L., Xia, X.C., He, Z.H., Gale, K.R., Lei, Z.S., Appels, R. and Ma, W. (2006) Characterization of three low-molecular-weight Glu-D3 subunit genes in common wheat. *Theor. Appl. Genet.* 113(7), 1247–1259.
- Zhao, X.L., Xia, X.C., He, Z.H., Lei, Z.S., Appels, R., Yang, Y., Sun, Q.X. and Ma, W. (2007) Novel DNA variations to characterize low molecular weight glutenin Glu-D3 genes and develop STS markers in common wheat. *Theor. Appl. Genet.* 114(3), 451–460.
- Zolla, L., Rinalducci, S., Timperio, A.M. and Huber, C. (2002) Proteomics of light-harvesting proteins in different plant species. Analysis and comparison by liquid chromatography-electrospray ionization mass spectrometry. *Photosystem I. Plant Physiol.* 130, 1938–1950.

Chapter 22

Linkage Disequilibrium and Association Mapping in the Triticeae

Mark E. Sorrells and Jianming Yu

Abstract Association mapping, also referred to as linkage disequilibrium mapping, has gained considerable popularity as an efficient genetic mapping methodology because of improved statistical approaches that increase power and reduce false positive associations. Association mapping exploits historical recombination events in a diverse population of distantly related or unrelated individuals. In this chapter, linkage disequilibrium estimates and association mapping results for the Triticeae are reviewed and compared to other species. Strategies for implementing association mapping are discussed that take into account objectives, types of markers, species and population size and composition. Finally, different methodologies for the application of association analysis to crop improvement are presented along with issues specific to breeding programs.

22.1 Introduction

The discipline of plant breeding plays a critical role in the utilization of biological information for the benefit of humankind. Because plant breeding is an integrative science, a multidisciplinary approach is required to successfully develop superior crop germplasm and varieties. New knowledge and technological advances contribute to crop improvement by increasing the efficiency of accurately selecting unique phenotypes and genotypes for target environments. Linkage mapping is fundamental to Mendelian genetics and molecular breeding. Molecular markers have greatly enhanced the resolution of genome mapping in those species without a complete genome sequence and contributed to our understanding of the genetic control of important traits. New methods and technologies that increase the efficiency of phenotyping, genome mapping and

M.E. Sorrells (✉)

Department of Plant Breeding and Genetics, Cornell University, Ithaca,
NY 14853, USA
e-mail: mes12@cornell.edu

data analysis contribute significantly to crop improvement. This review will present general concepts of linkage disequilibrium and association analyses and then discuss the strategies and approaches for effectively exploiting association analyses for crop improvement. While we have attempted to keep the background information general and balanced with respect to all crops, the review is focused on applications to members of the Triticeae, most of which are self-pollinated species. For other recent reviews the reader might consider Jannink et al. (2001), Gupta et al. (2005), Breseghello and Sorrells (2006a), Ersoz et al. (2007), MacKay and Powell (2007) or Yu and Buckler (2006).

22.2 Linkage Disequilibrium

Linkage disequilibrium (LD), or gametic disequilibrium, refers to the non-random association of alleles at linked or unlinked loci throughout the genome (Hedrick 1987) within a population and it forms the basis for association mapping (AM) strategies. Biparental cross populations have been widely used to determine the number and genome location of genes controlling traits of interest. Such populations have both advantages and limitations. In crosses between inbred, diploid or disomic parents, there can only be two alleles at each polymorphic locus. Because LD is maximized, fewer markers are required for genome coverage, but mapping resolution is minimal. Allele frequencies are usually close to 0.5 and there is no population substructure.

Knowledge of the level of genetic diversity and historical relationships among a collection of genotypes can be very useful for association mapping (AM) and the exploitation of genetic variation. The power of association analysis is affected by the patterns of LD, the extent of LD in the genome, and the variation in LD from one population to another. Linkage disequilibrium is affected by mating system, recombination rate, population structure, population history, genetic drift, directional selection and gene fixation (reviewed by Gaut and Long 2003). The utilization of ancestral recombination and the genetic diversity of a panel of distantly related genotypes are key components contributing to advantages of the AM approach compared to biparental cross populations. The number of generations since the occurrence of a common ancestor in the pedigree of any two genotypes in a population is an important determinant of the LD among loci, and consequently, the resolution of the mapping.

22.2.1 Measurement and Interpretation of Linkage Disequilibrium

The LD between alleles at two loci can be measured using quantitative metrics such as D , D' (Lewontin 1964), r^2 (Hill and Robertson 1968), or other statistics (e.g. Hedrick 1987; Pritchard and Przeworski 2001). $D_{ab} = P_{ab} - (p_a)(p_b)$,

where D_{ab} is a measure of LD between two loci, X and Y ; P_{ab} is the frequency of gamete X_aY_b ; p_a and p_b are the frequencies of the alleles X_a and Y_b at locus X and Y , respectively. These measures of LD largely differ in how they are affected by marginal allele frequencies (Hedrick 1987). The D' statistic is partially normalized with respect to allele frequencies. Another commonly used statistic is the squared value of the Pearson's (product moment) correlation coefficient, r^2 , that is a measure of the proportion of the variance of a response variable explained by a predictor variable (Hill and Robertson 1968). Because of the frequency dependence of r^2 , this statistic can be difficult to interpret when the frequencies of alleles at two loci are quite different. However, the intuitive nature of r^2 facilitates the interpretation of marker densities and association analyses. For example, if a causative DNA polymorphism, or quantitative trait nucleotide (QTN), is assumed to contribute a fraction of the total variation in a quantitative trait, we can estimate the fraction of the variance explained by a marker in LD. If the QTN has a heritability of h^2_Q (i.e., it explains that fraction of the phenotypic variance), then the fraction of the phenotypic variance explained by a marker in LD is $r^2 \times h^2_Q$. In addition, statistical power for detecting marker-phenotype associations is sometimes interpreted as a function of the parameter $4 N_e C$ where N_e is the effective population size and C is the recombination rate. For a large natural population without selection and mutation, Hill and Robertson (1968) deduced that $E[r^2] = 1/(4 N_e C + 1)$. For more detailed explanations of LD statistics see Hedrick (1987), Long and Langley (1999), Pritchard and Przeworski (2001), and Gaut and Long (2003).

At intermediate allele frequencies, D' and r^2 are correlated, but when allele frequencies are extreme, these two statistics can be quite different. One approach when using these statistics to measure LD between two loci with multiple alleles is to group related, low frequency alleles at a locus using a minimum allele frequency threshold (Weir 1996). The significance of association between any two loci can be determined using permutation (Long and Langley 1999) and can be displayed along with the LD statistic on opposite sides of a diagonal (e.g. Fig. 22.1). Such figures are useful for visualizing LD along a chromosome, or in the case of low LD, within a genic region.

Although LD has been reported to be highly variable within a genome and among different populations, the genome-wide LD is of interest as a general guide to marker density that may be required for whole-genome AM. Typically, association values for all pairwise linked or syntenic markers are plotted against either map distance or physical distance. Arbitrary values of $r^2 = 0.1$ – 0.2 are sometimes chosen as an indicator of statistically significant LD. This level of LD, however, would indicate that the closest marker only captures 20% of the phenotypic variation resulting from a causal polymorphism which is insufficient to generate detectable association between the marker and the phenotype. Brescghello and Sorrells (2006b) based whole-genome r^2 significance on unlinked estimates of r^2 . Those r^2 estimates were square-root transformed and the 95th percentile of that distribution was used as a threshold value.

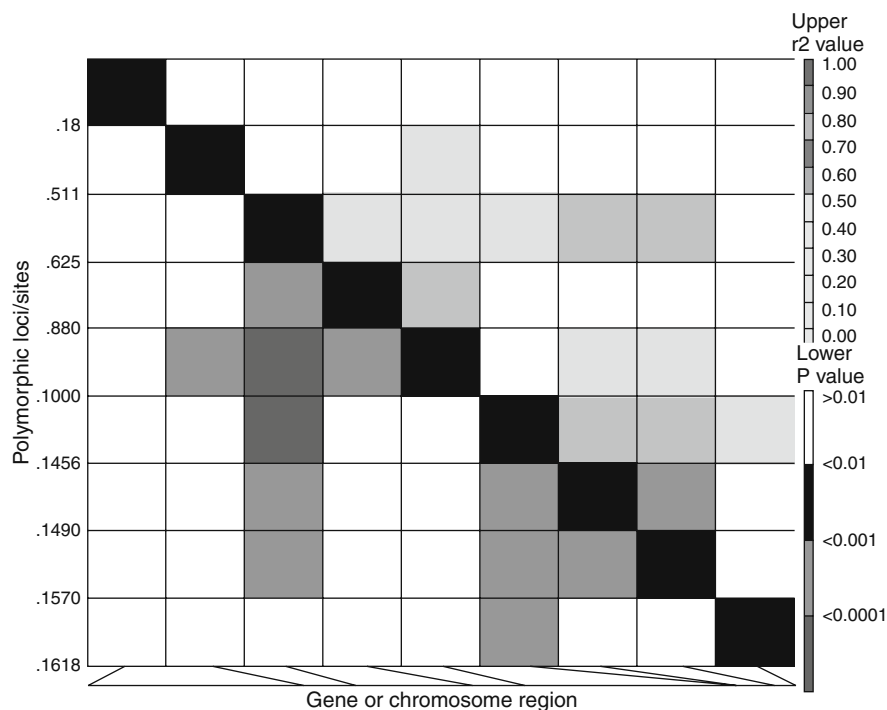


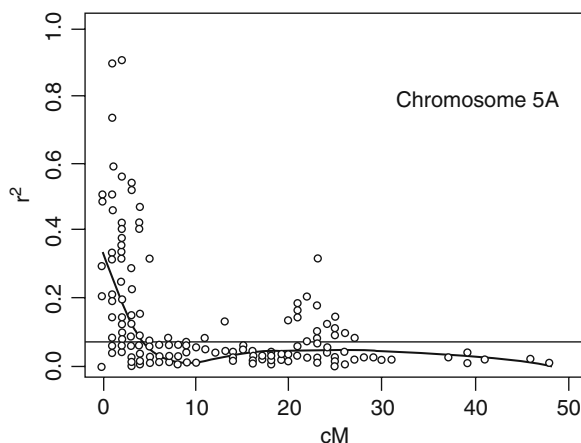
Fig. 22.1 Example of a figure showing pairwise r^2 values above the diagonal and corresponding p values below the diagonal

A second-degree loess curve was plotted through the data and the point of intersection with the threshold value was used as the genome-wide estimate of LD among loci (Fig. 22.2). This threshold value distinguishes between linked and unlinked loci, but higher LD (more closely linked markers) would be needed for AM.

22.2.2 LD Estimates for the Triticeae

Researchers often state that detailed knowledge of LD is necessary for AM. However, because LD is affected by many factors, there is much variation in the extent of LD within genomes, among different accessions within species, and among species (Remington et al. 2001; Flint-Garcia et al. 2003; Ersoz et al. 2007). Consequently, only a general idea of the level of LD for a species is possible and each population must be evaluated on a case-by-case basis. Nevertheless, it may still be possible to use sufficient marker density to identify blocks that exhibit only a few haplotypes in cultivated wheat and barley (Buntjer

Fig. 22.2 Estimates of r^2 vs. linkage distance on chromosome 5A. Horizontal straight lines indicate the 95th percentile of the distribution of unlinked r^2 . Curves were fit by second-degree loess. Adapted from Breseghello and Sorrells (2006b)



et al. 2005). For whole-genome scans, a haplotype map with diagnostic markers (also called tag SNPs) would be the most efficient methodology for AM (Gabriel et al. 2002).

One of the more prominent determinants of LD is the mode of reproduction, i.e. self-versus cross-pollination (Nordborg and Donnelly 1997). The reduced frequency of outcrossing in self-pollinated species limits chances for recombination. For example, estimates of LD in maize range from 200 bp (Tenaillon et al. 2001) to 2000 bp (Remington et al. 2001) versus 50 kbp in *Arabidopsis* (Nordborg et al. 2005) and centiMorgan distances in wheat (Chao et al. 2007; Maccaferri et al. 2006; Somers et al. 2007). However, wild relatives of self-pollinated species such as barley (Morrell et al. 2005; Steffenson et al. 2007) and ancestral populations (Caldwell et al. 2006; Maccaferri et al. 2006) can also show reduced LD. This is in part due to the fact that cultivated species have undergone strong selection pressure during the domestication process and distinct subpopulations, often related to geographic origin or adaptation, are evident (e.g. Dreisigacker et al. 2004; Melchinger et al. 1994).

Recent linkage disequilibrium studies for cultivated wheat and barley reported that LD estimates range from sub-centiMorgan up to 50 cM, much greater than reported for outcrossing species. There have been several assessments of LD in cultivated durum and hexaploid wheat, some of which have also examined LD in subpopulations (Maccaferri et al. 2005, 2006; Somers et al. 2007). This becomes important in germplasm panels that represent a wide range of genetic diversity (Yu et al. 2006) because some unlinked allele combinations are subpopulation specific. In two studies characterizing LD in cultivated durum wheat, Maccaferri et al. (2005, 2006) reported significant LD ($r^2 p > 0.1$) of 20 cM for 134 diverse worldwide accessions but only 10 cM for 189 Mediterranean cultivars that included a subset of the worldwide group. Long range LD was common in the worldwide collection with more than 50%

independent loci significantly associated compared to 23% within the Mediterranean group. Also, LD was lower in the Mediterranean group for all but the most closely linked class of loci. Somers et al. (2007) conducted a genome-wide evaluation of LD in 93 diverse durum and 189 bread wheat accessions. At $r^2 > 0.2$, LD extended only 2–3 cM for both species. In some genome regions, LD extended to 41 cM for bread wheat and 25 cM for durum. Within subpopulations, there were 14–16 fold fewer locus pairs in significant LD ($p < 0.001$) but the r^2 values were greater for both durum and bread wheat. Potentially, small subpopulation size was a contributing factor. Maccaferri et al. (2006) and Chao et al. (2007) came to similar conclusions regarding population size. The low genetic diversity reported for cultivated durum wheat may partially explain the high level of LD (Haudry et al. 2007).

Some hexaploid wheat studies have reported lower LD in cultivated accessions than that of the durum studies by Maccaferri et al. (2005, 2006). For example, Breseghello and Sorrells (2006b) reported significant LD ($r^2 > 0.065$) less than 1 cM and 5 cM for chromosomes 2D and 5A, respectively, in a panel of 95 elite, eastern soft winter wheat cultivars. Chao et al. (2007) assessed genetic relationships and LD over the entire genome of 43 diverse U.S. cultivars and breeding lines. Genome-wide LD estimates ($r^2 > 0.2$) were generally less than 10 cM although LD was highly variable throughout the genome. There was little evidence of LD around major genes such as those for vernalization or photoperiod sensitivity with the exception of a red kernel color locus on 3D. This was consistent with variation in LD reported by Breseghello and Sorrells (2006b) and known variation in recombination rates in the wheat genome (Akhunov et al. 2003). Somers et al. (2007) conducted a genome-wide assessment of LD in 189 hard red spring wheat accessions that mainly represented Canadian breeding programs. The LD estimates were 2–3 cM in general but extended to 41 cM in some regions. They also observed fewer haplotypes on group 3 chromosomes containing genes for red kernel color. When they plotted r^2 against chromosome position, there were regions of very high and low LD but no data were presented relating those regions to genes that may have been under selection. The authors concluded that LD mapping in their population could be accomplished using SSRs to achieve a resolution of less than 5 cM. Tommasini et al. (2007) examined LD on chromosome 3B in 44 European winter wheat varieties and found that r^2 extended less than 0.5 cM. In contrast, Crossa et al. (2007) reported that LD among DArT (Diversity Arrays Technology) markers declined within about 40 cM for two panels consisting of genotypes from five elite spring wheat disease nurseries. They calculated the average cM distance for 43 significant linkage blocks to be about 10 cM. The type of marker used can influence LD estimates. LD as measured by r^2 decreases as the number of alleles increase (Hedrick 1987). The r^2 values are inflated because population sizes are inadequate to assess the frequency of all haplotypes. This is less of a problem for markers that are biallelic. In the study by Crossa et al. (2007), only one DArT marker had an “absent” frequency below 5%. Because SSRs are multiallelic, it is not uncommon to observe rare alleles and those are

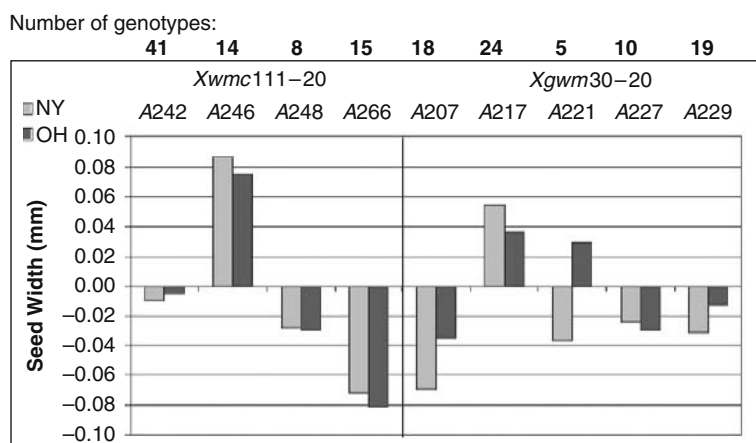


Fig. 22.3 Example of marker alleles associated with different QTL effects on the trait seed width. Adapted from Bresgheello and Sorrells (2006b)

typically excluded from LD estimates. The argument has been made that SNP markers are superior to SSR markers because of the multiallelic nature of SSRs. In some cases that may be true, however it may be more related to linkage distance than a result of more alleles. In practice, some SSR alleles have identical values with respect to the trait and can be lumped together for analysis (Fig. 22.3). In the case of haplotypes, the goal is to find new haplotypes that are linked to QTL that are superior to existing ones. As a result, we end up with multiple SNP haplotypes that need to be evaluated in the same way as SSRs.

For barley, LD estimates have varied widely. Using 236 AFLP markers and cultivated two-row barley varieties, Kraakman et al. (2004) reported that LD was at least 10 cM and varied widely over the genome. More recently, Rostoks et al. (2006) assayed 1,524 genome-wide SNP markers using 102 predominately northwestern European barley varieties along with a few parents of mapping populations. In the European spring barley, significant intrachromosomal LD extended over 60 cM, but in the spring, two-row subset only 15 cM. To validate the utility of this germplasm for AM, they re-mapped 140 previously mapped loci to establish a cutoff of $r^2 > 0.5$ (5% false positive rate). They then analyzed 85 unmapped loci and were able to place 43 loci to a putative map position. Those were compared to an orthologous region of the rice genome to predict the accuracy and 34 of 43 loci were validated. They concluded that larger population sizes would be required and that a denser map would increase the success rate of LD mapping. Malysheva-Otto et al. (2006) surveyed molecular diversity in a worldwide collection of 953 cultivated barley accessions. Using 48 SSR markers, they estimated LD at 10 cM for $r^2 < 0.2$. This population was highly structured according to the geographic origin.

Steffenson et al. (2007) estimated LD in a population of 318 wild barley (*Hordeum spontaneum*) accessions. LD estimates were quite low (<1 cM) for all chromosomes except 4H for which LD estimates extended up to 50 cM at $r^2 = 0.5$. Lin et al. (2002) also observed very low LD between closely linked *adh* loci in wild barley accessions. Strake et al. (2007) reported a detailed analysis of LD surrounding the barley gene conferring resistance to barley yellow mosaic virus. Haplotypes distinguished resistant from susceptible accessions in a population with diverse origin. They concluded that LD surrounding the resistance gene extended at least one centiMorgan, a distance that would be excessive for map-based gene cloning. Russell et al. (2003) evaluated genetic diversity and LD as measured by D' in a population of barley landrace varieties from different regions of the Fertile Crescent. They interpreted significant long-range LD as an indication that the adapted gene complexes and the assembly of interacting alleles into synergistic complexes were relevant to the evolution of landraces.

22.3 Association Analysis

22.3.1 Population Structure

To minimize statistical error, correction for population structure is critical in a collection of genotypes, especially in a breeding program where genetic relationships among breeding lines are highly variable. Essentially, this is due to the fact that genotypes in the collection for association analysis are rarely independent samples given the geographical origins, local adaptation, and breeding history of these genotypes (Yu and Buckler 2006). In other words, these genotypes do not constitute a designed mapping population comparable to the traditional F_2 , backcross, double haploid (DH), or recombinant inbred line (RIL) populations. One way to show the elevated false positive rate caused by population structure is to test the association between random SNPs and the trait of interest. Under the assumption of SNPs being random, one would expect roughly 5% of the tests to be significant, given a type I error rate of 5%. If a much higher proportion of SNPs turn out to be significant (i.e., P value ≥ 0.05), the model used for testing is determined to be inadequate and models considering population structure will need to be investigated.

However, the concern over false positives resulting from population structure has been alleviated by recent advances in methodology development. Notably, by using molecular markers that are randomly distributed across the genome, population structure of a diverse collection can be inferred and then taken into account in the further analysis of the candidate polymorphisms for functional association (Falush et al. 2003; Pritchard and Rosenberg 1999; Pritchard et al. 2000). For example, empirical studies of AM that consider population structure have been shown in maize (Thornsberry et al. 2001), wheat (Brescaghello and Sorrells 2006a), barley (Kraakman et al. 2004),

perennial ryegrass (Skøt et al. 2007), and *Arabidopsis* (Zhao et al. 2007). On another front, mixed models have been used for mapping genes with genotypes of plant breeding programs, for which known pedigree structure was used to account for the genetic relationship (Crepieux et al. 2005; Parisseaux and Bernardo 2004; Zhang et al. 2005).

A unified mixed model that simultaneously considers both population structure and familial relatedness was proposed to combine the strengths of the above approaches (Yu et al. 2006). This resulted from an understanding that most AM panels assembled in crops would contain some level of major subpopulation structure and some level of familial relatedness because of the nature of breeding and sampling. With this approach, false positives are controlled at the nominal level as has been demonstrated in maize and human (Yu et al. 2006), and in *Arabidopsis* (Zhao et al. 2007). It has been demonstrated that this combined approach, which is made possible by using background markers to determine population structure and relative kinship, is superior to either approach alone. Two steps prior to the mixed model analysis are the population structure analysis with STRUCTURE (Pritchard et al. 2000) and the relative kinship analysis with SPAGeDi (Hardy and Vekemans 2002). Model testing can be carried out to determine whether both population structure and relative kinship need to be included in the model. Empirical testing with random SNPs can also provide information on how well the false positive rate is controlled in the models (Yu et al. 2006; Zhao et al. 2007). Similarly, principal component analysis has also been proposed as a way to control for population structure (Price et al. 2006b), and was integrated into the unified mixed model framework (Weber et al. 2007; Zhao et al. 2007). In both cases, principal components replaced the subpopulation covariance matrix to account for the gross-level population differentiation but kinship remained in the unified mixed model. While the utility of the subpopulation covariance and kinship matrices are likely to vary from case to case, both studies further demonstrated the superiority of false positive control of the unified mixed model approach over other approaches.

22.3.2 Association Mapping Strategies

Association mapping strategies differ for different species as well as for different objectives and are still evolving as new methodologies are developed. In many species such as wheat, marker number and density are often the limiting factors, thus they require populations with high LD or strategies focusing on specific candidate genes or QTL intervals. For species that are not limited by marker resources, multiple testing and false discovery rate are more prominent technical issues. When fine mapping is the objective, the usual approach is to utilize unrelated (distantly related) germplasm and focus on specific regions of the genome believed to harbor a gene affecting the trait of interest. When QTL/gene discovery is the goal, whole genome scans in populations with high LD will be a

superior strategy due to the reduction in the number of markers required and the increased probability that a true association between genotype and phenotype will be found. That probability could be offset if the population with high LD is accompanied by reduced polymorphism for genes affecting the trait of interest. The types of markers can also affect the design of AM experiments. Although SSR markers have high information content, multiple alleles reduce the power and low frequency alleles inflate r^2 values in small populations. Biallelic markers simplify analyses but larger numbers are required to construct whole genome haplotypes and low allele frequencies reduce the information content. While there have been a few programs developed to facilitate AM, TASSEL has been the most commonly used program (Bradbury et al. 2007). TASSEL is a full-featured, cross-platform analysis tool that is frequently updated as new methods are developed.

Population size is one of the most influential factors determining the success of AM in detecting and estimating gene effects. Long and Langley (1999) examined the statistical power of AM to detect SNP/phenotype associations using simulations with varied quantitative trait nucleotide (QTN) effect, SNP density, and population size. As expected, increasing the number of SNPs and population size increased the power to detect associations. However, increasing population size had the most effect on power of the haploid marker permutation (HMP) test (Fig. 22.4). For QTN explaining 10% of the variation, power estimates were 0.2, 0.5, and 0.85 for population sizes of 50, 100, and 500 individuals, respectively. Even for a QTN accounting for 25% of the variation, power was only 0.6 and 0.8 for 50 and 100 individuals. For a population of 500 individuals, the power curve plateaued for QTNs accounting for over 10% of the variation. Sampling errors can also cause overestimation of gene effects

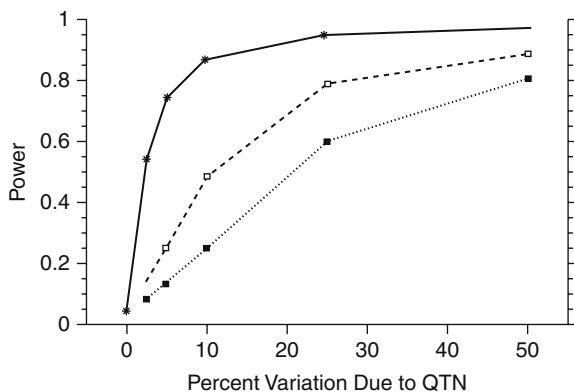


Fig. 22.4 Power of the haploid marker permutation test as a function of the simulated percentage of variation attributable to the quantitative trait nucleotide (QTN). *Dotted, dashed, and solid lines* are for experiment sizes of 50, 100, and 500 individuals, respectively (Reproduced from Long and Langley 1999)

because the error is confounded with the differences between alleles when the same data are used for detection and estimation of effects. Limited sampling of environments also contributes to error resulting from QTL \times environment interaction. Cross validation is important to confirm a locus effect and to estimate unbiased expectations of genetic gain. Similarly, cross validation has been proposed for linkage analysis with biparental cross populations by using independent data sets (Melchinger et al. 1998), preferentially from multiple environments, or by resampling from a larger data set (Schon et al. 2004).

Multi-stage AM is a strategy with the goal of reducing genotyping costs that is based on using a full set of markers on a small population followed by a larger population using a subset of selected markers that were significant in the first set based on a liberal statistical threshold (Hirschhorn and Daly 2005). Although there will be a certain percentage of false positive markers after the first screening, the second stage will use a second independent population that is similar or larger in size and the focused marker genotyping can result in a substantial increase in efficiency. Joint thresholds can be chosen for the first two stages so that the frequency of double-positives is quite low. Satagopan et al. (2007) investigated the efficiency of various multi-stage strategies for sequential QTL mapping. They concluded that an optimum approach for several scenarios was to genotype 60% of the individuals of a chosen population in stage 1 to identify the chromosomes significant at $\alpha = 0.20$ level and then use the remaining individuals in stage 2 for genotyping only those selected chromosomes. This sampling strategy was estimated to utilize only 70% of the genotyping burden relative to a single stage design regardless of heritability and marker density.

In contrast, the strategy of genotyping a larger population with a minimal number of markers allows the identification of a subset of accessions that may exhibit less population structure and fewer rare alleles (Breseghello and Sorrells 2006b). This can be especially important when the relationships among accessions are not known *a priori* and nearly identical genotypes are selected by chance. This multi-stage approach was used to “normalize” a population consisting of elite soft winter wheat varieties whose pedigrees were largely unknown (Fig. 22.5). Although initial genotyping costs may be higher, for studies requiring a large number of markers such as whole-genome scans, this method could actually be more cost effective than multi-stage approaches that use a subset of markers on a larger population. However, the relative efficiencies will depend on the cost of phenotyping and the degree of population structure. Also, the use of a kinship covariance matrix in the mixed linear model may offset some of the effects of the closely related genotypes (Yu et al. 2006).

22.3.3 Association Mapping in the Triticeae

Wheat has the fewest molecular markers available among the major crop species including barley, maize, rice, sorghum, and soybean. Consequently,

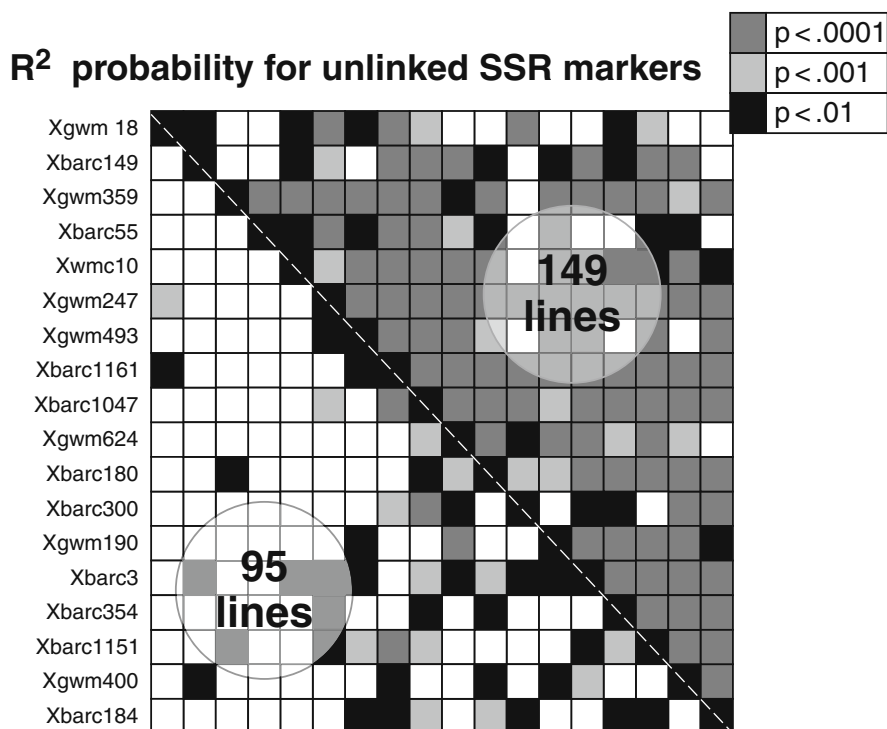


Fig. 22.5 r^2 significance values for unlinked loci using 149 elite soft winter wheat varieties and the selected set of 95 that were used by Breseghello and Sorrells (2006b) for association mapping

whole genome scans to date have been relatively low resolution. For wheat, LD and AM studies have mostly employed SSR markers that are subject to the limitations noted above. Low resolution LD studies using SSRs were reported by Chao et al. (2007), Maccaferri et al. (2005, 2006), and Somers et al. (2007) but AM of traits was not included in those reports. The recent development of wheat DArT markers (Wenzl et al. 2004; Akbari et al. 2006) has facilitated whole genome analyses. Crossa et al. (2007) used DArT markers to identify associations with stem, leaf, and stripe rust resistance in historical international wheat trials. Using mixed linear models, they identified loci associated with rust resistance in regions of the genome previously known to harbor resistance genes as well as new chromosome regions. They also modeled $G \times E$ and additive \times additive effects of markers on traits. Targeted wheat AM studies were reported by Ravel et al. (2006), Breseghello and Sorrells (2006b) and Tommasini et al. (2007). Breseghello and Sorrells (2006b) conducted AM using a mixed effects model on 95 selected eastern soft winter wheat varieties for milling quality, seed size and seed shape. Their analysis was focused on two previous QTL intervals

mapped in unrelated, biparental cross populations. Those QTL effects were validated and fine mapped in the AM panel. In the study by Tommasini et al. (2007), 44 modern European winter wheat varieties were used to validate and fine map the location of a QTL for *Stagonospora nodorum* blotch resistance that was previously mapped in a biparental cross population. The Sun2-3B marker was confirmed in these varieties demonstrating that the resistance allele is conserved in current breeding germplasm. Ravel et al. (2006) used association study to successfully discriminate between two candidate genes controlling a QTL for the quantity of high molecular weight glutenin subunit. These studies suggest that AM studies can be especially useful for fine mapping in wheat because of the time savings, relevance to breeding germplasm, and limited polymorphism in this species.

The availability of more marker resources has benefited barley researchers. An early study using 236 AFLP markers examined associations between markers and yield and yield components in a collection of 146 European two-row spring barley cultivars (Kraakman et al. 2004). The marker/trait associations confirmed in this association panel agreed with the locations of QTL for yield reported in earlier QTL mapping studies. Those reports used distantly related North American germplasm some of which had been used as parents of biparental cross populations. They suggested that the confirmed QTL may have had their origin in common ancestors of the two gene pools. The authors also used multiple linear regression to select 18–20 loci explaining 40–58% of the variation in yield. Steffenson et al. (2007) used association analysis to map genes controlling leaf, stripe, and stem rust in 318 wild barley (*Hordeum spontaneum*) accessions. A wide range of infection types and resistance was observed for all three rusts (including the stem rust race TTKS; i.e. Ug99). DArT markers were used for AM in the wild barley population revealing 14 or 15 marker associations (8 or 9 bins) depending on the number of subpopulations used. A major association with stem rust on 5H was validated using a biparental cross population.

22.3.4 Germplasm Panels

Germplasm suited for AM can be broadly classified into three categories: exotic accessions from germplasm bank collections, intermated populations, and elite lines (Brescghello and Sorrells 2006a). These classes of germplasm can be used for different purposes according to their genetic expectations. Those expectations are conditioned by the number of meiotic events since a common ancestor between any two accessions and their genetic relationship (Remington et al. 2001). However, there is little information available regarding criteria that could be used to assemble germplasm panels to maximize the information for addressing specific objectives. Accessions have been chosen to represent the maximum amount of genetic diversity in a species (e.g. Pestova and Röder 2002;

Maccaferri et al. 2005), a regional gene pool (Maccaferri et al. 2006; Somers et al. 2007), elite breeding germplasm (Brescghello and Sorrells 2006b), or natural or wild populations (Nordborg et al. 2002; Morrell et al. 2005). In general the goal is to maximize the amount of information generated with limited resources. For some traits that are strongly affected by environment and subject to $G \times E$, there are tradeoffs between the level of genetic diversity among the accessions in the panel and the ability to accurately evaluate the trait. This is because some complex traits such as biomass yield are strongly affected by general adaptation. A simple example would be evaluating grain yield of winter wheat planted in the spring. Clearly, one or a few genes are affecting the adaptation and preventing the expression of genes that may be important for grain yield if it were grown in a region where winter wheat would be vernalized.

Brescghello and Sorrells (2006a) elaborated the practical aspects, genetic expectations, and applications of these three categories of germplasm (Table 22.1). A core collection of accessions obtained from a germplasm bank may be used to screen high heritability traits, whereas elite lines in a breeding program are usually evaluated for low heritability traits in replicated, multi-environment trials. Progenies from an intermated population can be evaluated in different ways depending on the recurrent selection method and the traits. The genetic expectations for an exotic core collection are low LD, low to medium population structure, and high allelic diversity. Intermated populations synthesized from inbreds start out with high LD which is eroded with additional cycles of intermating. Elite lines have high LD and population structure resulting from the recent crossing and selection of elite, adapted lines. Consequently, AM power is low and resolution is high for exotic germplasm, while expectations are the opposite for elite lines. Intermated populations synthesized from inbreds start out with high power and low resolution, both of which are moderated through cycles of intermating. Breeders use these contrasting sources of germplasm for different purposes. Exotic germplasm is typically used as a source of novel alleles in a marker-assisted backcross scheme whereas elite lines are intermated and marker-assisted selection is used in the segregating progenies in a forward-breeding strategy. Intermated segregating populations offer a favorable balance of power and precision for association analysis and would allow mapping of quantitative traits with increasing resolution through cycles of intermating. The question of how to select germplasm for an association study has not been addressed in a comprehensive analysis for plants. Human studies have many constraints that do not exist for plants. Also, plant diversity, even within a species generally exceeds that of *Homo sapiens*. The availability of germplasm collections, complex mating designs, large populations, and the lack of ethical issues greatly facilitate genetics studies in plants. Based on results of association analysis studies in plants it is clear that the makeup of the population of accessions chosen for association analysis has major effects on hypothesis testing and accomplishing stated objectives. Further research to evaluate different strategies for selection of germplasm for different goals is needed.

Table 22.1 Comparison of different types of populations for association analysis

Aspects of association analysis	Germplasm bank	Elite germplasm	Intermating populations
Samples	Core collection	Elite lines and cultivars	Individual plants or progenies
Sample turnover	Static	Gradually substituted	Ephemeral
Source of phenotypic data	Screenings	Replicated yield trials	Progeny testing
Type of traits	High heritability and domestication traits	Low heritability traits: yield, abiotic stress tolerance	Depends on the evaluation scheme
Level of LD	Low	High	Intermediate and fast decaying
Population structure	Medium	High	Low
Allele diversity among samples	High	Low	Intermediate
Allele diversity within samples	Variable [†]	1 allele	1 or 2 alleles [‡]
Use of information from unlinked loci	Study organization of genetic variation	Variety protection; control of seed purity	Monitor allele diversity, population structure and effective size
Resolution of AA	High	Low	Intermediate and increasing
Power of AA	Low	High	Intermediate and decreasing
Use of significant markers	Marker-assisted backcross	Forward breeding, genomic selection	Incorporation in selection index

[†] Depends on the collection, conservation, and sampling schemes.

[‡] For diploid species.

22.3.5 Simulations

Computer simulations have been extensively used to investigate various factors in associate analysis, such as sample size, marker density, trait heritability, genetic relationships, false positive control, and population types (Arbelbide et al. 2006; Yu et al. 2005, 2006, 2009). The usefulness of computer simulation resides in the ability to investigate a wide range of scenarios with different genetic and non-genetic parameters, rather than a specific case study in an empirical association analysis. However, insights gained from empirical analyses are critical for setting up these computer simulations because they provide

realistic parameters and bounds. They also frequently highlight the importance of particular issues that arise in practical applications.

There is still limited knowledge about the effect of population size on detection power in AM in crop species. Using the data from the maize diversity panel, it was found that an average power of 0.8 was obtained for a QTL explaining 3.3% of phenotypic variation in flowering time when the full set of 274 maize inbreds was used, whereas the same power was only possible for a QTL explaining more than 10% of the phenotypic variation when the sample size was only 69 maize inbreds were used (Yu, unpublished data). In general, a moderate sample size (greater than 137 maize lines) increased the statistical power to detect causative polymorphisms with relatively small to moderate genetic effects explaining 1–12% of the phenotypic variation in a trait. Detection of moderate genetic effects is likely to be important for many quantitative traits in plants and animals (Yu, unpublished data). One thing that needs to be clarified is that the percentage of the phenotypic variation explained by QTL in AM is generally much smaller than that in linkage analysis because of the much homogenized genetic context in the latter case. Recent genome-wide AM in *Arabidopsis* using 95 diverse accessions also raised the issue of sample size for identifying functional polymorphisms with moderate to small effects (Zhao et al. 2007).

The number of background markers required for accurate estimates of population structure and kinship is a common issue that needs to be addressed in many AM studies. For the unified mixed model, both population structure and relative kinship are estimated based on molecular markers. A likelihood based model fitting approach was recently proposed to quantify the robustness of genetic relationships derived from molecular marker data (Yu et al. 2007b). In theory, the adequacy of markers in relationship estimation (i.e. population structure and relative kinship) influences the maximum likelihood of the model explaining phenotypic variation. In other words, a more accurate relationship matrix based on more markers provides a better quantification of genetic variance-covariance underlying the phenotypic variation. This was shown to be the case with computer simulations under various scenarios. Using empirical data from maize and canine, it was demonstrated that kinship construction with subsets of the whole marker panel and subsequent model testing with multiple phenotypic traits can provide *ad hoc* information on whether the number of markers is sufficient to quantify genetic relationships among individuals. On the other hand, population structure was less sensitive to the number of markers as shown by model fitting and variance component estimation (Yu et al. 2009). In general, the number of multiallelic SSRs required for genetic relationship estimation is smaller than that of biallelic SNPs. In the case of the data from maize, a species with high nucleotide diversity, 89 SSRs provided the same level of relationship quantification as 912 SNPs based on model fitting criteria (Yu et al. 2009). Association panels with complex population structure and familial relatedness would require more markers than those with simpler structure. In practice, about

100 SSRs may be adequate for most crop association panels and the adequacy can be tested with multiple trait data as outlined in Yu et al. (2009).

A series of computer simulations have also been conducted to investigate the usefulness of two breeding schemes in maize: Marker Assisted Recurrent Selection (MARS) and Genome-Wide Selection (GWS). Genome-wide selection or genomic selection was first reported for animal breeding and involves estimating allelic effects for all markers and all markers are used for predicting breeding value (Meuwissen et al. 2001). Both GWS and MARS breeding methods were designed to integrate QTL mapping and marker assisted selection into regular breeding programs and are discussed below under “Breeding Applications”. Further theoretical and empirical works are needed to implement MARS and/or GWS in self-pollinated species such as Triticeae species, for which random mating of selected individuals is very labor intensive and seeds available for initial field testing are usually limited. The principals underlying MARS and GWS, however, should apply equally to both self- and cross-pollinated species.

22.4 Future Needs and Directions

22.4.1 *Fine-Mapping*

Higher mapping resolution has been regarded as one of the merits for association analysis as compared with the traditional linkage analysis using biparental cross populations. Also a broader inference space of genetic effects would be advantageous in marker assisted breeding. The average level of resolution depends on the genome-wide LD decay, which in turn depends on the species and the type of population examined, and may vary in different regions of the genome (Brescghello and Sorrells 2006b; Flint-Garcia et al. 2003; Somers et al. 2007). The classical map-based cloning approach has so far successfully cloned or tagged 12 major effect QTL and 9 small effect QTL (Price 2006a). In addition to the huge cost associated with these studies, the time scale from initial QTL mapping to final cloning or tagging is usually several years. Association mapping has the potential to achieve the fine mapping of functional polymorphisms underlying various complex traits within a much shorter time after the establishment and initial analysis of the panel. In fact, recent QTL cloning/tagging studies have utilized the haplotype structure surrounding the targeted genome region across a diverse set of accessions as supporting evidence of the candidate gene (Li et al. 2006). Although no formal association analyses were conducted, the basic principle was the same.

Genome-wide studies with high density SNPs may eventually capture the full potential of AM as exemplified in studies searching for causal mutations of several human diseases (WTCCC 2007). Community-wide resources can also be important for building an information base (Churchill et al. 2004). However, initial scans with moderate SNP coverage in Arabidopsis revealed both opportunities and challenges (Zhao et al. 2007). Many issues need to be addressed, including population size, genetic relationships, SNP coverage, phenotyping

precision, and multiple testing corrections, before whole genome scans with diverse collections of accessions and breeding populations become practical for the Triticeae.

On the other hand, joint linkage and linkage disequilibrium mapping have been proposed as a fine mapping approach in theory (Mott and Flint 2002; Wu and Zeng 2001; Wu et al. 2002), and demonstrated in practice (Blott et al. 2003; Meuwissen et al. 2002). Nested AM, as is currently implemented in maize, seems to be an even more powerful strategy for dissecting the genetic basis of complex quantitative traits in species with low LD (Yu et al. 2008). Nested AM involves selecting a set of diverse founders that are crossed to a common parent to generate a large set of related mapping progenies. By integrating genetic design, natural diversity, and genomics technologies, the nested AM strategy allows a high-power, fine resolution, cost-effective genome scan and should facilitate endeavors to link molecular variation with phenotypic variation for various complex traits. Because intermating in self-pollinated species such as barley and wheat is quite labor intensive, alternative approaches are needed to reduce LD in these species. While LD may be low enough in the wild relatives for fine mapping, many agronomic traits cannot be evaluated in wild relatives, thus preventing the association of functional polymorphism with phenotype. Mackay and Powell (2007) have advocated the development of multiparent advanced generation intercross populations originally proposed by Darvasi and Soller (1995) and evaluated by Mott et al. (2002). In principle, this approach has some of the same advantages as NAM. Several founders representing a significant portion of the variation in a target population are intermated for multiple generations and then a large population of inbred lines is derived by selfing. This approach is labor intensive during the intermating stage but inbred lines of self-pollinated crops are easily produced. Also, it accomplishes the important objective of reducing LD in a cultivated, self-pollinated species which is usually a limiting factor. Another approach is the use of populations segregating for male-sterility. Several methods of population improvement for self-pollinated species have been elaborated using either a recessive (Brim and Stuber 1973) or a dominant male-sterile gene (Sorrells and Fritz 1982). However, such populations are potentially useful resources for fine mapping because of the enforced recombination nearly every generation (Heffner et al. 2008). Other methods of reducing LD in the Triticeae are needed to facilitate the fine mapping of candidate genes that affect traits unique to these crop species.

22.4.2 Breeding Applications

Association analysis and complex trait dissection can be integrated into conventional breeding programs using molecular tools and information to facilitate marker-assisted selection of parents and segregating populations. Breeding progress is derived from the (i) discovery and generation of genetic

variation for agronomic traits, (ii) development of genotypes with new or improved attributes due to superior combinations of alleles at multiple loci, and (iii) accurate selection of rare genotypes that possess the new improved characteristics. Breeding programs are dynamic, complex genetic entities in which there is a constant flux of alleles moving in and out of the gene pool. Consequently, a molecular breeding strategy requires frequent evaluation of marker/phenotype relationships. Jansen et al. (2003) described an extension of the multiple QTL models developed by Jannink and Jansen (2001) for multiple related populations. Their method was based on dense marker maps and conserved haplotype sharing of genotypes in a breeding program to map QTL.

Because biparental cross populations sometimes involve poorly adapted parents, exhibit maximum linkage disequilibrium and are limited to two alleles per locus, the application of information to a breeding program can be quite limited. Association mapping can be conducted directly on the breeding material greatly facilitating the practical use of the information in a crop improvement program. Because there is more genetic diversity in a breeding program than in a biparental cross, phenotypic variation and marker polymorphism are usually much higher. Genotypic data can be combined with phenotypic data from routine screening and variety trial evaluations to facilitate selection for low heritability traits. Probably the most important advantage for a breeding program is that novel alleles can be identified and the relative allelic value can be assessed as often as necessary. In this discussion, we will address three approaches: (1) association breeding: crossing/selection/testing program, (2) marker-assisted recurrent selection (MARS), and (3) genome-wide selection.

22.4.3 Association Breeding

When considering the incorporation of association analysis in a breeding program there are several important issues to consider. Because marker-based selection involves selecting individuals possessing a marker allele that is associated with an allele at a locus controlling a phenotype of interest, the association between those alleles can be described as a conditional probability (Brescaghello and Sorrells 2006a). The estimation of gene effects using molecular markers is susceptible to errors resulting from sampling variance and systematic biases. As discussed earlier, overestimation of gene effects resulting from sampling bias is caused by the error being confounded with the differences between alleles when the same data set is used for both detection and estimation of effects. Also, population structure results in increased false positive associations as previously mentioned. In a typical breeding program, there is likely to be a wide range in genetic relationships among the genotypes as well as heterogeneity. This can be effectively dealt with using covariates for population structure and kinship (Yu et al. 2006) in a mixed linear model. Another approach described earlier is to select a subset of genotypes that have reduced population structure.

The complexity of the genetic constitution of a breeding program presents some novel issues for association analysis. Probably the most serious is the potential presence of marker allele/trait allele relationships that are different from that expected from previous studies (Brescghello and Sorrells 2006a). The case of an exclusive marker allele closely linked to the desired trait allele maximizes the conditional probability, however, should the same marker allele be present in the breeding population but linked to a neutral or undesirable trait allele, the association is drastically reduced. In fact, the presence of such an undesirable linkage introduced into the population at a frequency of only 5% reduces the marker/trait allele association more than if there were 10% recombination between the marker and trait loci over the first 8 generations of intermating, after which close linkage predominately affects the association (Fig. 22.6) (Brescghello and Sorrells 2006a). In fact, a short time frame for selection and intermating or inbreeding is characteristic of most breeding programs. The short time frame also influences the gain from selection of a marker allele that is loosely linked to the trait allele (Fig. 22.7). Under intense selection pressure, the opportunities for recombination between the marker and the trait alleles are minimized, thus the gain from selection, even for markers that are loosely linked, can be reasonably high. Once the marker/trait alleles are homozygous there is no chance for recombination. Therefore, the short time frame for selection and intermating or inbreeding can have a significant impact on the outcome of marker-assisted breeding.

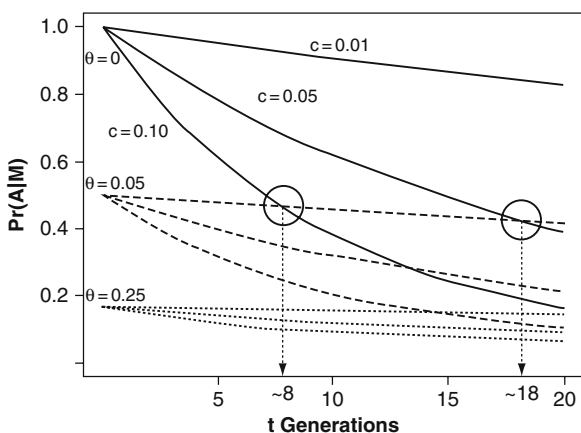
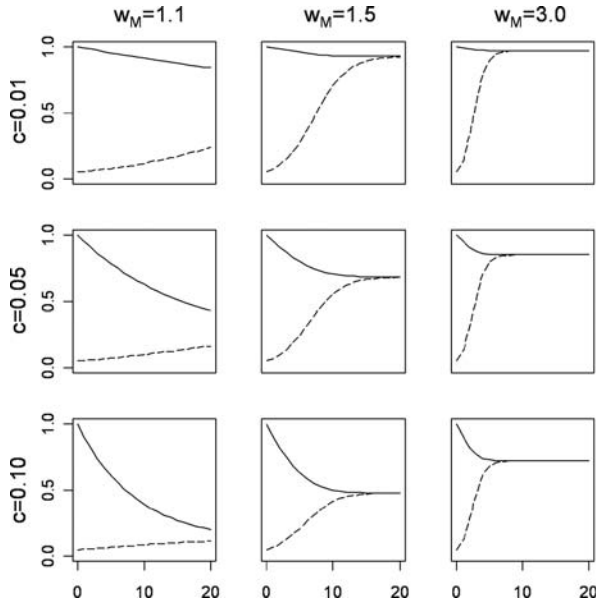


Fig. 22.6 Effect of a pre-existing marker allele in a breeding population illustrating that a novel marker allele at 10 cM distance can be more predictive of the QTL allele than an allele 1 cM away if it was already present in the original pop at a frequency of 0.05. $\Pr(A|M)$ = Conditional probability of a QTL allele A, given marker allele M. Frequency of marker allele in original population = θ ; Recombination = c

Fig. 22.7 Effect of selection intensity on frequency of QTL allele in a breeding population over generations for recombination frequencies (c) of 0.01, 0.05, and 0.10. Relative fitness = w_m



In a typical breeding program, the starting point is a source of genetic variation (Fig. 22.8). Selected genotypes are crossed to produce new populations that are subject to phenotypic and/or genotypic selection. Those materials are either intermated or inbred to produce new populations or inbred lines that are evaluated in replicated, multi-environment trials. Some of those trial entries are eventually released as new varieties. However, within the breeding program,

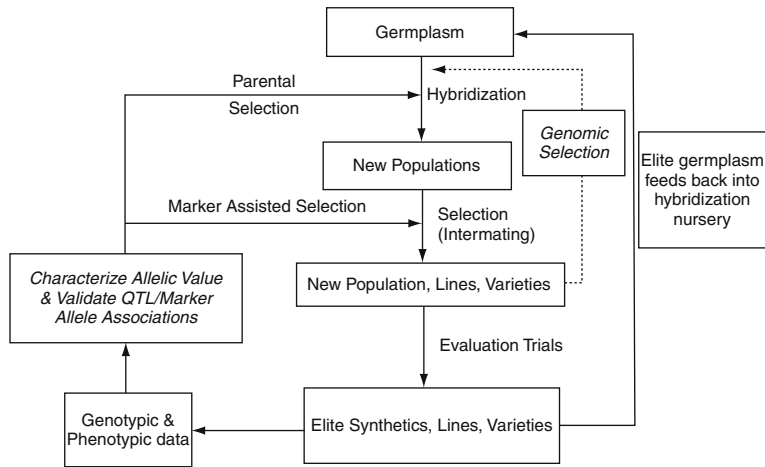


Fig. 22.8 Application of association analysis and genomic selection in a breeding program

the breeder uses the trial information to select elite parents that re-enter the hybridization program. In association breeding, genotypic data (preferably whole genome coverage) and the appropriate analyses are incorporated to validate previously mapped marker/trait associations and potentially identify new ones. The information is used to estimate allelic value at selected loci (or all loci in genomic selection) and then create a genotypic value index for each genotype and trait (Lande and Thompson 1990; Christopher et al. 2007). That information is used to select parents for hybridization and segregating or inbred individuals in the breeding nurseries. Most breeding programs consist of highly structured populations resulting from selection of different numbers of inbreds from crosses among moderately divergent parental lines (Fig. 22.8) that can result in a high frequency of false positive marker/trait associations. Further, low polymorphism may limit the utility of many markers and subsequent analyses. Other unforeseen issues are likely as breeders attempt to substitute genotypic for phenotypic information.

22.4.4 Marker Assisted Recurrent Selection

Marker assisted recurrent selection involves the improvement of an F_2 or F_1 -derived doubled haploid (DH) population using one generation of phenotypic and genotypic evaluation to identify marker/trait associations followed by multiple cycles of recurrent selection using only allelic values at the selected marker loci (Johnson 2001, 2004; Bernardo and Yu 2007). Various factors affecting the genetic gain of MARS over phenotypic selection have been studied through computer simulation in maize, including initial mapping population size for establishing marker-trait relationship, relaxed significance levels (e.g., $\alpha = 0.1$ or 0.2 rather than 0.01), marker numbers, and selection intensity (Bernardo and Charcosset 2006; Bernardo et al. 2006). Charmet et al. (1999) evaluated effects of population size, number of QTL and accuracy of QTL detection on recurrent selection strategy with an example using dough strength in wheat (Charmet et al. 2001). They used an algorithm to identify optimum crosses for combining all positive QTL alleles in a recurrent crossing and selection scheme. Two cycles of selection in a population of 200 was as efficient as a single cycle using 4,000. Christopher et al. (2007) used selected SSR markers to follow alleles in wheat and barley breeding programs. Changes in allele frequencies over generations at loci known to be associated with quality or disease resistance were consistent with expectations. Graphical genotypes were useful for tracing markers through generations. Peleman and Van der Voort (2003) elaborated a strategy referred to as “Breeding by Design” that proposed to identify allelic variation at all loci affecting agronomically important characters and design superior genotypes *in silico*.

22.4.5 Genomic Selection

The availability of inexpensive molecular markers for some crops has changed our point of view as to how markers may be used in crop improvement. The question has been posed “How can molecular markers best be used to achieve breeding progress” (Bernardo and Yu 2007). Genomic selection (alias genome-wide selection (Bernardo and Yu 2007)) involves marker-assisted selection in which selections are based on all markers across the genome rather than just those showing significant effects. Individuals in a phenotyped population are genotyped for markers distributed over the genome and then breeding values of all the markers are fitted as random effects in a linear model. Individuals in subsequent recurrent selection generations are then selected based on the sum of those breeding values (Meuwissen et al. 2001) (Fig. 22.8). Bernardo and Yu (2007) compared MARS to genome-wide selection in simulations involving a population of 144 individuals from which 4 individuals were selected in cycles 1 and 2. For MARS, a selection index based on all selected markers was calculated as suggested by Lande and Thompson (1990). For GWS, the best linear unbiased predictor (BLUP) of breeding values was estimated by fitting all the markers as random effects and imposing a convenient assumption of equal variances based on cycle 0 performance. Genome-wide selection resulted in a larger response to selection than MARS. Depending on the number of QTL and the heritability, the response to genome-wide selection was 6–18% higher than MARS with the biggest advantage for complex traits with low heritability. When the population was small and heritability low, selection of only the major QTL led to greater response to selection using MARS. For both selection methods, the cycle 0 population size was important with evaluation of a minimum of 96 DHs providing the maximum response to genotypic selection. The authors concluded by describing genome-wide selection as a black-box procedure where large numbers of markers are used as a surrogate for phenotype and it does not involve dissecting or understanding of the underlying physiologic or genetic mechanisms.

More efficient methods will continue to be developed for identifying and evaluating allelic effects on a large scale so that breeders can assemble desirable alleles in superior varieties. As we expand our knowledge of how genes evolve and interact to produce the nearly infinite range of phenotypes, new opportunities to manipulate genetic variation to the benefit of humankind will arise.

References

- Akbari, M., Wenzl, P., Caig, V., Carling, J., Xia, L., Yang, S., Uszynski, G., Mohler, V., Lehmensiek, A. and Kuchel, H. (2006) Diversity arrays technology (DArT) for high-throughput profiling of the hexaploid wheat genome. *TAG Theor. Appl. Genet.* 113, 1409–1420.
- Akhunov, E. D., Goodyear, A. W., Geng, S., Qi, L. L., Echaliier, B., Gill, B. S., Gustafson, J. P., Lazo, G., Chao, S. and Anderson, O. D. (2003) The organization and rate of evolution

- of wheat genomes are correlated with recombination rates along chromosome arms. *Genome Res.* 13, 753–763.
- Arbelbide, M., Yu, J. and Bernardo, R. (2006) Power of mixed-model QTL mapping from phenotypic, pedigree, and marker data in self-pollinated crops. *Theor. Appl. Genet.* 112, 876–884.
- Bernardo, R. and Charcosset, A. (2006) Usefulness of gene information in marker-assisted recurrent selection: a simulation appraisal. *Crop Sci.* 46, 614–621.
- Bernardo, R., Moreau, L. and Charcosset, A. (2006) Number and fitness of selected individuals in marker-assisted and phenotypic recurrent selection. *Crop Sci.* 46, 1972–1980.
- Bernardo, R. and Yu, J. (2007) Prospects for genomewide selection for quantitative traits in maize. *Crop Sci.* 47, 1082–1090.
- Blott, S., Kim, J. J., Moisisio, S., Schmidt-Kuntzel, A., Cornet, A., Berzi, P., Cambisano, N., Ford, C., Grisart, B., Johnson, D., Karim, L., Simon, P., Snell, R., Spelman, R., Wong, J., Vilkkki, J., Georges, M., Farnir, F. and Coppieters, W. (2003) Molecular dissection of a quantitative trait locus: a phenylalanine-to-tyrosine substitution in the transmembrane domain of the bovine growth hormone receptor is associated with a major effect on milk yield and composition. *Genetics* 163, 253–266.
- Bradbury, P., Zhang, Z., Kroon, D., Casstevens, T., Ram-Doss, Y. and Buckler, E. (2007) TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* 23, 2633–2635.
- Breseghele, F. and Sorrells, M. E. (2006a) Association analysis as a strategy for improvement of quantitative traits in plants. *Crop Sci.* 46, 1323–1330.
- Breseghele, F. and Sorrells, M. E. (2006b) Association mapping of kernel size and milling quality in wheat (*Triticum aestivum* L.) cultivars. *Genetics* 172, 1165–1177.
- Brim, C. A. and Stuber, C. W. (1973) Application of genetic male sterility to recurrent selection schemes in soybeans. *Crop Sci.* 14, 528–530.
- Buntjer, J., Sørensen, A. and Peleman, J. (2005) Haplotype diversity: the link between statistical and biological association. *Trends Plant Sci.* 10, 466–471.
- Caldwell, K., Russell, J., Langridge, P. and Powell, W. (2006) Extreme population-dependent linkage disequilibrium detected in an inbreeding plant species, *Hordeum vulgare*. *Genetics* 172, 557–567.
- Chao, S., Zhang, W., Dubcovsky, J. and Sorrells, M. E. (2007) Evaluation of genetic diversity and genome-wide linkage disequilibrium among wheat (*Triticum aestivum* L.) germplasm representing different market classes. *Crop Sci.* 47, 1018–1030.
- Charmet, G., Robert, N., Perretant, M. R., Gay, G., Sourdille, P., Groos, C., Bernard, S. and Bernard, M. (1999) Marker-assisted recurrent selection for cumulating additive and interactive QTLs in recombinant inbred lines. *Theor. Appl. Genet.* 99, 1143–1148.
- Charmet, G., Robert, N., Perretant, M. R., Gay, G., Sourdille, P., Groos, C., Bernard, S. and Bernard, M. (2001) Marker assisted recurrent selection for cumulating QTLs for bread-making related traits. *Euphytica* 119, 89–93.
- Christopher, M., Mace, E., Jordan, D., Rodgers, D., McGowan, P., Delacy, I., Banks, P., Sheppard, J., Butler, D. and Poulsen, D. (2007) Applications of pedigree-based genome mapping in wheat and barley breeding programs. *Euphytica* 154, 307–316.
- Churchill, G. A., Airey, D. C., Allayee, H., Angel, J. M., Attie, A. D., Beatty, J., Beavis, W. D., Belknap, J. K., Bennett, B., Berrettini, W., Bleich, A., Bogue, M., Broman, K. W., Buck, K. J., Buckler, E., Burmeister, M., Chesler, E. J., Cheverud, J. M., Clapote, S., Cook, M. N., Cox, R. D., Crabbe, J. C., Crusio, W. E., Darvasi, A., Deschepper, C. F., Doerge, R. W., Farber, C. R., Forejt, J., Gaile, D., Garlow, S. J., Geiger, H., Gershenfeld, H., Gordon, T., Gu, J., Gu, W., de Haan, G., Hayes, N. L., Heller, C., Himmelbauer, H., Hitzemann, R., Hunter, K., Hsu, H. C., Iraqi, F. A., Ivandic, B., Jacob, H. J., Jansen, R. C., Jepsen, K. J., Johnson, D. K., Johnson, T. E., Kempermann, G., Kendziorski, C., Kotb, M., Kooy, R. F., Llamas, B., Lammert, F., Lassalle, J. M., Lowenstein, P. R., Lu, L., Lusi, A., Manly, K. F., Marcucio, R., Matthews, D., Medrano, J. F., Miller, D. R., Mittleman, G., Mock, B. A.,

- Mogil, J. S., Montagutelli, X., Morahan, G., Morris, D. G., Mott, R., Nadeau, J. H., Nagase, H., Nowakowski, R. S., O'Hara, B. F., Osadchuk, A. V., Page, G. P., Paigen, B., Paigen, K., Palmer, A. A., Pan, H. J., Peltonen-Palotie, L., Peirce, J., Pomp, D., Pravenec, M., Prows, D. R., Qi, Z., Reeves, R. H., Roder, J., Rosen, G. D., Schadt, E. E., Schalkwyk, L. C., Seltzer, Z., Shimomura, K., Shou, S., Sillanpaa, M. J., Siracusa, L. D., Snoeck, H. W., Spearow, J. L., Svenson, K. et al. (2004) The collaborative cross, a community resource for the genetic analysis of complex traits. *Nat. Genet.* 36, 1133–1137.
- Crepieux, S., Lebreton, C., Flament, P. and Charmet, G. (2005) Application of a new IBD-based QTL mapping method to common wheat breeding population: analysis of kernel hardness and dough strength. *Theor. Appl. Genet.* 111, 1409–1419.
- Crossa, J., Burgueno, J., Dreisigacker, S., Vargas, M., Herrera-Foessel, S., Lillemo, M., Singh, R., Trethowan, R., Warburton, M., Franco, J., Reynolds, M., Crouch, J. and Ortiz, R. (2007) Association analysis of historical bread wheat germplasm using additive genetic covariance of relatives and population structure. *Genetics* 177, 1889–1913.
- Darvasi, A. and Soller, M. (1995) Advanced intercross lines, an experimental population for fine genetic mapping. *Genetics* 141, 1199–1207.
- Dreisigacker, S., Zhang, P., Warburton, M. L., Van Ginkel, M., Hoisington, D., Bohn, M. and Melchinger, A. E. (2004) SSR and pedigree analyses of genetic diversity among CIMMYT wheat lines targeted to different megaenvironments. *Crop Sci.* 44, 381.
- Ersoz, E. S., Yu, J. and Buckler, E. (2007) Applications of linkage disequilibrium and association mapping in crop plants. In: R. Varshney and R. Tuberosa (Eds.), *Genomics-Assisted Crop Improvement*. Springer, Dordrecht, Netherlands.
- Falush, D., Stephens, M. and Pritchard, J. K. (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* 164, 1567–1587.
- Flint-Garcia, S., Thornsberry, J. and Buckler, E. (2003) Structure of linkage disequilibrium in plants. *Annu. Rev. Plant Biol.* 54, 357–374.
- Gabriel, S. B., Schaffner, S. F., Nguyen, H., Moore, J. M., Roy, J., Blumenstiel, B., Higgins, J., DeFelice, M., Lochner, A. and Faggart, M. (2002) The structure of haplotype blocks in the human genome. *Science* 296, 2225–2229.
- Gaut, B. S. and Long, A. D. (2003) The lowdown on linkage disequilibrium. *Am. Soc. Plant Biol.* 15, 1502–1506.
- Gupta, P., Rustgi, S. and Kulwal, P. (2005) Linkage disequilibrium and association studies in higher plants: present status and future prospects. *Plant Mol. Biol.* 57, 461–485.
- Hardy, O. J. and Vekemans, X. (2002) SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Mol. Eco. Notes* 2, 618–620.
- Haudry, A., Cenci, A., Ravel, C., Bataillon, T., Brunel, D., Poncet, C., Hochu, I., Poirier, S., Santoni, S., Glémin S. and David J. (2007) Grinding up wheat: A massive loss of nucleotide diversity since domestication. *Mol. Biol. Evol.* 24, 1506–1517.
- Hedrick, P. W. (1987). Gametic Disequilibrium Measures: Proceed With Caution. *Genetics* 117, 331–341.
- Heffner, E. L., Ornumbol, C., Williams, K. R. and Sorrells, M. E. (2008) Dominant male-sterile populations for association mapping and introgression of exotic wheat germplasm. *Aust. J. Agric. Res.* 59, 470–474.
- Hill, W. G. and Robertson, A. (1968) Linkage disequilibrium in finite populations. *TAG Theor. Appl. Genetic.* 38, 226–231.
- Hirschhorn, J. N. and Daly, M. J. (2005) Genome-wide association studies for common diseases and complex traits. *Nat. Rev. Genet.* 6, 95–108.
- Jannink, J. L., Bink, M. and Jansen, R. C. (2001) Using complex plant pedigrees to map valuable genes. *Trends Plant Sci.* 6, 337–342.

- Jannink, J. L. and Jansen, R. (2001) Mapping epistatic quantitative trait loci with one-dimensional genome searches. *Genetics* 157, 445–454.
- Jansen, R. C., Jannink, J. L. and Beavis, W. D. (2003) Mapping quantitative trait loci in plant breeding populations: use of parental haplotype sharing. *Crop Sci.* 43, 829–834.
- Johnson, L. (2001) Marker assisted sweet corn breeding: a model for specialty crops. In: *Proceedings of the 56th Annual Corn and Sorghum Industrial Research Conference*, Chicago, IL. 5–7 Dec. 2001. American Seed Trade Association, Washington, DC, pp. 25–30.
- Johnson, R. (2004). Marker-assisted selection. *Plant Breed. Rev.* 24(1), 293–309.
- Kraakman, A. T. W., Niks, R. E., Van den Berg, P., Stam, P. and Van Eeuwijk, F. A. (2004) Linkage disequilibrium mapping of yield and yield stability in modern spring barley cultivars. *Genetics* 168, 435–446.
- Lande, R. and Thompson, R. (1990) Efficiency of marker-assisted selection in the improvement of quantitative traits. *Genetics* 124, 743–756.
- Lewontin, R. C. (1964) The interaction of selection and linkage. II. Optimum models. *Genetics* 50, 757–782.
- Li, C., Zhou, A. and Sang, T. (2006) Rice domestication by reducing shattering. *Science* 311, 1936–1939.
- Lin, J. Z., Morrell, P. L. and Clegg, M. T. (2002) The influence of linkage and inbreeding on patterns of nucleotide sequence diversity at duplicate alcohol dehydrogenase loci in wild barley (*Hordeum vulgare* ssp. *spontaneum*). *Genetics* 162, 2007–2015.
- Long, A. D. and Langley, C. H. (1999) The power of association studies to detect the contribution of candidate genetic loci to variation in complex traits. *Genome Res.* 9, 720–731. Cold Spring Harbor Lab.
- Maccaferri, M., Sanguineti, M. C., Natoli, V., Ortega, J. L. A., Salem, M. B., Bort, J., Chenenaoui, C., De Ambrogio, E., del Moral, L. G. and De Montis, A. (2006) A panel of elite accessions of durum wheat (*Triticum durum* Desf.) suitable for association mapping studies. *Plant Genet. Resour.: Characterization and Utilization* 4, 79–85.
- Maccaferri, M., Sanguineti, M. C., Noli, E. and Tuberosa, R. (2005) Population structure and long-range linkage disequilibrium in a durum wheat elite collection. *Mol. Breeding* 15, 271–290.
- Mackay, I. and Powell, W. (2007) Methods for linkage disequilibrium mapping in crops. *Trends Plant Sci.* 12, 57–63.
- Malysheva-Otto, L. V., Ganal, M. W. and Röder, M. S. (2006) Analysis of molecular diversity, population structure and linkage disequilibrium in a worldwide survey of cultivated barley germplasm (*Hordeum vulgare* L.). *BMC Genetics* 7, 6.
- Melchinger, A. E., Graner, A., Singh, M. and Messmer, M. M. (1994) Relationships among European barley germplasm. I: Genetic diversity among winter and spring cultivars revealed by RFLPs. *Crop Sci.* 34, 1191–1199.
- Melchinger, A. E., Utz, H. F. and Schon, C. C. (1998) Quantitative trait locus (QTL) mapping using different testers and independent population samples in maize reveals low power of QTL detection and large bias in estimates of QTL effects. *Genetics* 149, 383–403.
- Meuwissen, T. H., Hayes, B. J. and Goddard, M. E. (2001) Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157, 1819–1829.
- Meuwissen, T. H., Karlsen, A., Lien, S., Olsaker, I. and Goddard, M. E. (2002) Fine mapping of a quantitative trait locus for twinning rate using combined linkage and linkage disequilibrium mapping. *Genetics* 161, 373–379.
- Morrell, P. L., Toleno, D. M., Lundy, K. E. and Clegg, M. T. (2005) Low levels of linkage disequilibrium in wild barley (*Hordeum vulgare* ssp. *spontaneum*) despite high rates of self-fertilization. *Proc. National Acad. Sci.* 102, 2442–2447.
- Mott, R., Talbot, C. J., Turri, M. G., Collins, A. C. and Flint, J. (2002) A method for fine mapping quantitative trait loci in outbred animal stocks. *Proc. Nat. Acad. Sci. USA* 97, 12649–12654.

- Mott, R. and Flint, J. (2002) Simultaneous detection and fine mapping of quantitative trait loci in mice using heterogeneous stocks. *Genetics* 160, 1609–1618.
- Nordborg, M., Borevitz, J. O., Bergelson, J., Berry, C. C., Chory, J., Hagenblad, J., Kreitman, M., Maloof, J. N., Noyes, T., Oefner, P. J., Stahl, E. A. and Weigel, D. (2002) The extent of linkage disequilibrium in *Arabidopsis thaliana*. *Nat. Genet.* 30, 190–193.
- Nordborg, M. and Donnelly, P. (1997) The coalescent process with selfing. *Genetics* 146, 1185–1195.
- Nordborg, M., Hu, T. T., Ishino, Y., Jhaveri, J., Toomajian, C., Zheng, H., Bakker, E., Calabrese, P., Gladstone, J., Goyal, R., Jakobsson, M., Kim, S., Morozov, Y., Padhukasaahasram, B., Plagnol, V., Rosenberg, N. A., Shah, C., Wall, J. D., Wang, J., Zhao, K., Kalbfleisch, T., Schultz, V., Kreitman, M. and Bergelson J. (2005) The pattern of polymorphism in *Arabidopsis thaliana*. *PLoS Biol.* 3:e196.
- Parisseaux, B. and Bernardo, R. (2004) In silico mapping of quantitative trait loci in maize. *Theor. Appl. Genet.* 109, 508–514.
- Peleman, J. D. and van der Voort, J. R. (2003) Breeding by design. *Trends Plant Sci.* 8, 330–334.
- Pestsova, E. and Röder, M. (2002) Microsatellite analysis of wheat chromosome 2D allows the reconstruction of chromosomal inheritance in pedigrees of breeding programmes. *Theor. Appl. Genet.* 106, 84–91.
- Price, A. H. (2006a) Believe it or not, QTLs are accurate! *Trends Plant Sci.* 11, 213–216.
- Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. A. and Reich, D. (2006b) Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* 38, 904–909.
- Pritchard, J. K. and Przeworski, M. (2001) Linkage disequilibrium in humans: models and data. *Am. J. Human Genet.* 69, 1–14.
- Pritchard, J. K. and Rosenberg, N. A. (1999) Use of unlinked genetic markers to detect population stratification in association studies. *Am. J. Human Genet.* 65, 220–228.
- Pritchard, J. K., Stephens, M. and Donnelly, P. (2000) Inference of population structure using multilocus genotype data. *Genetics* 155, 945–959.
- Ravel, C., Praud, S., Murigneux, A., Linossier, L., Dardevet, M., Balfourier, F., Dufour, P., Brunel, D. and Charmet, G. (2006) Identification of *Glu-B1-I* as a candidate gene for the quantity of high-molecular-weight glutenin in bread wheat (*Triticum aestivum* L.) by means of an association study. *Theor. Appl. Genet.* 112, 738–743.
- Remington, D. L., Thornsberry, J. M., Matsuoka, Y., Wilson, L. M., Whitt, S. R., Doebley, J., Kresovich, S., Goodman, M. M. and Buckler, E. S. (2001) Structure of linkage disequilibrium and phenotypic associations in the maize genome. *Proc. Natl. Acad. Sci. USA* 98, 11479–11484.
- Rostoks, N., Ramsay, L., MacKenzie, K., Cardle, L., Bhat, P., Roose, M., Svensson, J., Stein, N., Varshney, R., Marshall, D., Graner, A., Close, T. and Waugh, R. (2006) Recent history of artificial outcrossing facilitates whole-genome association mapping in elite inbred crop varieties. *Proc. Natl. Acad. Sci. USA* 103, 18656–18661.
- Russell, J. R., Booth, A., Fuller, J. D., Baum, M., Ceccarelli, S., Grando, S. and Powell, W. (2003) Patterns of polymorphism detected in the chloroplast and nuclear genomes of barley landraces sampled from Syria and Jordan. *TAG Theor. Appl. Genet.* 107, 413–421.
- Satagopan, J. M., Sen, S. and Churchill, G. A. (2007) Sequential quantitative trait locus mapping in experimental crosses. *Stat. Appl. Genet. Mol. Biol.* 6, Article12.
- Schon, C. C., Utz, H. F., Groh, S., Truberg, B., Openshaw, S. and Melchinger, A. E. (2004) Quantitative trait locus mapping based on resampling in a vast maize testcross experiment and its relevance to quantitative genetics for complex traits. *Genetics* 167, 485–498.
- Skot, L., Humphreys, J., Humphreys, M. O., Thorogood, D., Gallagher, J., Sanderson, R., Armstead, I. P. and Thomas, I. D. (2007) Association of candidate genes with

- flowering time and water-soluble carbohydrate content in *Lolium perenne* (L.). *Genetics* 177, 535.
- Somers, D., Banks, T., Depauw, R., Fox, S., Clarke, J., Pozniak, C. and McCartney, C. (2007). Genome-wide linkage disequilibrium analysis in bread wheat and durum wheat. *Genome* 50, 557–567.
- Sorrells, M. E. and Fritz, S. E. (1982) Application of a dominant male-sterile allele to the improvement of self-pollinated crops. *Crop Sci.* 22, 1033–1035.
- Steffenson, B. J., Olivera, P., Roy, J. K., Jin, Y., Smith, K. P. and Muehlbauer, G. J. (2007) A walk on the wild side: mining wild wheat and barley collections for rust resistance genes. *Austl. J. Agric. Res.* 58, 532.
- Strake, S., Prestler, T., Stein, N., Perovic, D., Ordon, F. and Graner, A. (2007) Effects of introgression and recombination on haplotype structure and linkage disequilibrium surrounding a locus encoding Bymovirus resistance in barley. *Genetics* 175, 805–817.
- Tenaillon, M.I., Mark C. Sawkins, M. C., Long, A.D., Gaut, R. L., Doebley, J. F. and Gaut, B. S. (2001) Patterns of DNA sequence polymorphism along chromosome 1 of maize (*Zea mays* ssp. *mays* L.). *Proc. Natl. Acad. Sci. USA* 98, 9161–9166.
- Thornsberry, J. M., Goodman, M. M., Doebley, J., Kresovich, S., Nielsen, D. and Buckler, E. S. (2001) Dwarf8 polymorphisms associate with variation in flowering time. *Nat. Genet.* 28, 286–289.
- Tommasini, L., Schnurbusch, T., Fossati, D., Mascher, F. and Keller, B. (2007) Association mapping of *Stagonospora nodorum* blotch resistance in modern European winter wheat varieties. *Theor. Appl. Genet.* 115, 697–708.
- Weber, A., Clark, R., Vaughn, L., de Jesús Sánchez-Gonzalez, J., Yu, J., Yandell, B. S., Bradbury, P. and Doebley, J. (2007) Major regulatory genes in maize contribute to standing variation in teosinte (*Zea mays* ssp. *parviglumis*). *Genetics* 177, 2349–2359.
- Weir, B. S. (1996) *Genetic Data Analysis II: Methods for Discrete Population Genetic Data*. Sinauer Associates Inc., USA.
- Wenzl, P., Carling, J., Kudrna, D., Jaccoud, D., Huttner, E., Kleinohs, A. and Kilian, A. (2004) Diversity arrays technology (DART) for whole-genome profiling of barley. *Proc. Natl. Acad. Sci. USA* 101, 9915–9920.
- WTCCC. (2007) Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 447, 661–678.
- Wu, R., Ma, C. X. and Casella, G. (2002) Joint linkage and linkage disequilibrium mapping of quantitative trait loci in natural populations. *Genetics* 160, 779–792.
- Wu, R. and Zeng, Z. B. (2001) Joint linkage and linkage disequilibrium mapping in natural populations. *Genetics* 157, 899–909.
- Yu, J., Arbelbide, M. and Bernardo, R. (2005) Power of in silico QTL mapping from phenotypic, pedigree, and marker data in a hybrid breeding program. *Theor. Appl. Genet.* 110, 1061–1067.
- Yu, J. and Buckler, E. (2006) Genetic association mapping and genome organization of maize. *Curr. Opin. Biotechnol.* 17, 155–160.
- Yu, J., Holland, J. B., McMullen, M. D. and Buckler, E. S. (2008) Genetic design and statistical power of nested association mapping in maize. *Genetics* 178, 539–551.
- Yu, J., Pressoir, G., Briggs, W., Vroh Bi, I., Yamasaki, M., Doebley, J., McMullen, M., Gaut, B., Nielsen, D., Holland, J., Kresovich, S. and Buckler, E. (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat. Genet.* 38, 203–208.
- Yu, J., Zhang, Z., Zhu, C., Tabanao, D., Pressoir, G., Tuinstra, M. R., Kresovich, S., Todhunter, R. S. and Buckler, E. S. (2009). Simulation appraisal of the adequacy of number of background markers for relationship estimation in association mapping. *The Plant Genome* 2, xxx–xxx (in press).

- Zhang, Y. M., Mao, Y., Xie, C., Smith, H., Luo, L. and Xu, S. (2005) Mapping quantitative trait loci using naturally occurring genetic variance among commercial inbred lines of maize (*Zea mays* L.). *Genetics* 169, 2267–2275.
- Zhao, K., Aranzana, M. J., Kim, S., Lister, C., Shindo, C., Tang, C., Toomajian, C., Zheng, H., Dean, C., Marjoram, P. and Nordborg, M. (2007) An *Arabidopsis* example of association mapping in structured samples. *PLoS Genet.* 4, 71–82.

Chapter 23

Triticeae Genome Structure and Evolution

Jan Dvořák

Abstract Repeated nucleotide sequences are by far the largest component of Triticeae genomes, accounting for about 90% of the nuclear DNA. Tandem repeated sequences play an important role in chromosome function during mitosis and meiosis. Interspersed repeated sequences fill the intergenic regions. The most remarkable attribute of this component is its unprecedented rate of turnover, which is in stark contrast to the stability of gene content. The term “gene order paradox” is coined to reflect this dichotomy. A model is proposed postulating the existence of two strata in Triticeae genomes, “conservative” and “dynamic” to account for this paradox and its evolutionary causes. Numerous aspects of gene content in Triticeae genomes, such as the location of single-copy genes, multi-gene loci, gene deletions and duplications, gene density, restriction fragment length polymorphism and single nucleotide polymorphism, location of novel and lineage-specific genes, and the level of synteny, correlate with recombination rate and gene location on the centromere-telomere axis. Special attention is devoted to the discussion of gene distribution along chromosomes. It is pointed out that evidence for the existence of gene-rich islands is weak. A model accounting for correlation between gene density and recombination rate is proposed. It is suggested that the vast amounts of repeated sequences in Triticeae genomes play a role in the evolution of new genes and in adaptation.

23.1 Structure of Triticeae Genomes

23.1.1 Genome Size

Löve (1984) listed in his Triticeae monograph 313 species, of which 18% were diploid, with chromosome number $2x = 14$ and 45.9% were polyploid. The chromosome number for the rest was unknown at the time of the publication.

J. Dvořák (✉)

Department of Plant Sciences, University of California, Davis, CA, USA
e-mail: jdvorak@ucdavis.edu

DNA content has been estimated for many species (The Kew Royal Botanical Garden *C*-value Database). Neglecting the disproportionately low values of two species in the database which need reinvestigation, the smallest genome sizes within the tribe were recorded in diploid genera *Critesion*, $1C = 3,455$ and $3,305$ Mb in *C. pusillum* ssp. *pusillum* and ssp. *euclaston*, respectively, (Jakob et al 2004) and *Pseudoroegneria*, $1C = 3,896$ Mb in *P. libanotica* and $3,920$ Mb in *P. stipifolia* (Vogel et al. 1999). The genome size of the annual diploid *Aegilops tauschii* is slightly larger, $4,024$ Mb (Arumuganathan and Earle 1991). Diploid species of genera *Psathyrostachys*, $1C = 7,644$ – $8,771$ Mb (Vogel et al. 1999), and *Secale*, $1C = 7,081$ – $8,110$ Mb (Bennett and Smith 1976), lie at the other end of the genome size spectrum in the tribe.

Because genome sizes have already been estimated for more than one-third of all Triticeae species, it seems safe to conclude that a large genome size is the norm across the tribe and may actually be its universal feature. Triticeae genomes lead in this respect in the entire grass family. *Psathyrostachys* and *Secale* are at the top among all the diploid grass species in the Kew Royal Botanical Garden *C*-value Database, being exceeded only by *Briza media* and *B. maxima* (Bennett and Smith 1976), and the largest genome size in the grass family has been reported for decaploid *Lophopyrum* (syn. *Thinopyrum*) *ponticum*, $1C = 22,173$ Mb (Vogel et al. 1999). Shedding light on the significance of the vast amounts of DNA in the Triticeae genomes is one of the great challenges facing Triticeae genetics and genomics.

23.1.2 Overall Structure

Classical DNA reassociation kinetics studies showed that repeated nucleotide sequences are the largest component of Triticeae genomes. For example, genomes of *Hordeum vulgare*, *Triticum monococcum*, and *Secale cereale*, representative species of three different Triticeae clades, were shown to contain 76, 80, and 92% repeated sequences, respectively (Flavell et al. 1974). In wheat, 4–10% of DNA reassociated instantly, presumably representing DNA that reassociated intramolecularly, and 70–80% reassociated with kinetics characteristic of a repetition frequency ranging from 100 to 100,000 copies per hexaploid genome. Only about 12% of wheat DNA reassociated with kinetics of single-copy DNA (Smith and Flavell 1975). Repeated sequences in the Triticeae genomes fall into two basic classes: interspersed sequences and tandem repeated sequences (satellites).

Sequencing of random clones or ends of large-insert clones, such as bacterial artificial chromosome (BAC) clones, was more recently used to assess the overall structure of Triticeae genomes (Li et al. 2004; Akhunov et al. 2005; Devos et al. 2005; Sabot et al. 2005; Paux et al. 2006). Comparing these sequences with repeated sequence databases, such as the Triticeae Repeated Sequence Database (TREP; <http://wheat.pw.usda.gov/ITMI/Repeats/>), which contains Triticeae repeated elements, and RepBase (<http://www.girinst.org/repbase/update/index.html>), which contains eukaryotic repeated elements, estimates the portion of genomic DNA

sequence represented by known repeated sequences. Comparison of BAC-end sequences of different Triticeae genomes in a single study and against the same version of a repeated sequence database showed that different genomic components vary in a narrow range in different species indicating that the overall genome structure of Triticeae genomes is similar (Akhunov et al. 2005).

Known and unknown repeated sequences accounted for 85.9% of BAC-end sequences of wheat chromosome 3B (Paux et al. 2006). In the *Ae. tauschii* random plasmid library, known and unknown repeated sequences accounted for 91.6% of the total DNA sequence (Li et al. 2004). Li et al. (2004) estimated unknown repeats by dot blot hybridization, which indicates that a clone contains a repeated sequence but does not estimate its proportion in the total sequence of a clone. Since some repeats may be embedded in nonrepeated DNA, this estimate could somewhat overestimate the actual proportion of unknown repeated sequences, and the two estimates may actually be closer to each other than is indicated by their face values.

The BAC-end sequence approach in hexaploid wheat estimated the tandem-repeated fraction to be 0.1% of the total chromosome 3B DNA (Paux et al. 2006), whereas the random cloning and sequencing approach estimated this fraction in *Ae. tauschii* to be 2.6% of the total DNA (Li et al. 2004). Because the BAC library was constructed by partial digestion of chromosome 3B DNA with the *HindIII* endonuclease (Šafář et al. 2004), and because the *HindIII* restriction site may be underrepresented in some tandem-repeated arrays, the former approach very likely underestimated the tandem-repeated fraction. Considering these estimates in their totality, tandem and interspersed repeated sequences may account for as much as 90% or even slightly more, of the total DNA in Triticeae genomes.

Coding sequences accounted for 1.2% of the chromosome 3B BAC-end sequences (Paux et al. 2006) and 2.5% of the *Ae. tauschii* random clone sequences (Li et al. 2004). In addition, 5.9% of the *Ae. tauschii* sequences were single-copy noncoding sequences (Li et al. 2004). Together, nonrepeated nucleotide sequences, presumably genes and surrounding *cis*-regulatory elements, accounted for 8.4% of the total *Ae. tauschii* DNA. In another attempt to estimate the percentage of DNA representing genes, 3.8 Mb of sequenced BAC DNA was analyzed (Sabot et al. 2005). Genes accounted for 7.8% of the total DNA length, yielding a similar estimate as the *Ae. tauschii* study.

It appears that an average Triticeae genome may contain about 8% genes and gene related sequences and 90% repeated nucleotide sequences. In light of the fact that repeated DNA represents such a large component of the Triticeae genomes, the structure and evolutionary dynamics of these genomes can hardly be understood without understanding the structure and evolution of repeated sequences.

23.1.3 Tandem Repeated Sequences

Most tandem repeats are located in the centromeric and telomeric regions of Triticeae chromosomes, but some also exist at interstitial sites (Fig. 23.1).

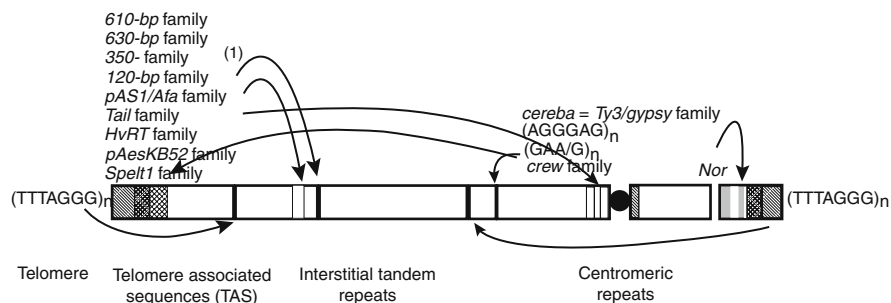


Fig. 23.1 Schematic summary of the locations and translocations of tandem repeated sequence families and rDNA in Triticeae chromosomes. Arrows indicate translocations of sequences and refer only to the families or sites from which they emanate

Tandem repeats are usually associated with constitutive heterochromatin (henceforth heterochromatin) in the Triticeae genomes and can be visualized by the C-banding procedure (Gill and Kimber 1974) (see Chapter 4). If present in a sub-critical amount, tandem repeats may fail to produce a C-band. Such sites can be visualized by fluorescent *in situ* DNA hybridization (FISH). An estimate of the total length of tandem repeated sequences in Triticeae genomes can be obtained by summing up the lengths of C-bands. In the A genome of Chinese Spring wheat (*T. aestivum*), the sum of the length of C-bands accounts for about 4% of the total chromosome length, but in the B genome, it accounts for as much as 15% of the total chromosome length (J. Dvořák, unpublished). In the genome of *Secale cereale*, C-bands in the subtelomeric regions account for 12.2% of the total chromosome length (Bennett et al. 1977).

23.1.3.1 Centromeric Regions

The centromeric regions of the Triticeae chromosomes consist of the primary constriction and pericentromeric heterochromatin. Complete and truncated copies of *Ty3/gypsy* retroelement-like sequences, first isolated from a sorghum centromere (Miller et al. 1998), were shown to exist in the primary constrictions of barley, wheat, *Aegilops*, and rye chromosomes (Presting et al. 1998; Fukui et al. 2001; Cheng and Murata 2003). A complete element was isolated from a barley primary constriction and named *cereba* (Presting et al. 1998). The element has the landmark features of retroelements, i.e., the retroviral poly-protein gene including the integrase and long terminal repeats (LTRs) nearly 1 kb long (Hudakova et al. 2001). LTRs were postulated to contain the *cereba* centromeric sequence (*CCS1*) (Presting et al. 1998), earlier described by Aragon-Alcaide et al. (1996), but the equivalence of these two sequences has been questioned (Hudakova et al. 2001). It was estimated that there are an average of 70 copies of the *cereba* sequences per *Ae. tauschii* centromere (Fukui et al. 2001). In hexaploid wheat, repeated copies of the retroelement are present in all 21 centromeres with an average spacing of about 55 kb between them (Fukui

et al. 2001). In barley, *cereba* elements and a *G + C* satellite sequence consisting of an AGGGAG motif represent the core of the barley centromeres (Hudakova et al. 2001). The *G + C* satellite was not detected in wheat centromeres, although the sequence appears to be related to the wheat GA(A/G) polypurine/polypyrimidine satellite (see below). In wheat, the tandem repeat *crew*, a 250 bp long sequence presumably derived from a *Ty3/gypsy* element, is interspersed with *cereba* units (Cheng and Murata 2003).

Sequence motifs of centromeric repeats are usually distinct from those present in the subtelomeric regions. However, some repeat families, exemplified by the *TaiI* family, may be present in subtelomeric regions in one species and centromeric regions in another Triticeae species (Kishii et al. 2001).

Appels et al. (1978) described a satellite sequence consisting of polypurine/polypyrimidine tracts analogous to the *Drosophila melanogaster* 1.705 g/cm³ satellite sequence. This satellite was later shown to consist of a simple trinucleotide sequence GA(A/G) (Dennis et al. 1980). While in rye the sequence is at several interstitial sites and represents a mere 0.1% of total DNA, it is in the pericentromeric regions of the seven wheat B-genome chromosomes and accounts for about 3% of the total hexaploid wheat genomic DNA (Dennis et al. 1980; Peacock et al. 1981). This sequence is the principal component of the centromeric heterochromatin in the B genome but is also present in subtelomeric sites of several chromosomes (Gerlach et al. 1979; Peacock et al. 1981). The location of GA(A/G) polypurine/polypyrimidine repeats coincides with the location of N-bands (Gerlach et al. 1979; Pedersen et al. 1996; Cuadrado et al. 2000), although the reverse is not always true (Pedersen et al. 1996; Cuadrado and Jouve 2007). Tracts of polypurine/polypyrimidine sequences were detected in centromeric, interstitial, and subtelomeric sites in the genomes of *Aegilops* species of section *Sitopsis*, barley, *Elymus*, and *Hordeum* (Dennis et al. 1980; Pedersen et al. 1996).

Whereas the presence of tandem repeats in the centromeric regions is conserved in Triticeae and other plants, the primary sequence of the repeats and their organization is not and varies even among very closely related species (Kishii et al. 2001; Amor et al. 2004). It was suggested that tandem repeats are a structural DNA theme for the formation of heterochromatin which has important epigenetic functions for the centromere (Grewal and Moazed 2003; Amor et al. 2004). Pericentromeric heterochromatin plays a critical role in sister chromatid adhesion and proper orientation of sister centromeres towards the poles (Bernard et al. 2001). Modification of H3 and H4 histones creates epigenetic memory, via which the position of kinetochore is stably transmitted from generation to generation (Grewal and Moazed 2003).

This model accounts for otherwise conflicting findings obtained with Triticeae centromeres. The *cereba* repeats are conserved across Triticeae and grasses. They interact with heterochromatin-specific histone H3 (CENH3) indicating that they are a functionally important component of cereal centromeres (Houben et al. 2007). Yet, deletion of all *cereba* repeats and the *G + C* satellite sequence from a barley chromosome had no effect on the regularity of

its mitotic and meiotic behavior. Translocation of these repeats onto wheat chromosome 4A did not generate a dicentric chromosome (Nasuda et al. 2005). At their face value, these observations seem to suggest that these repeats are dispensable and not sufficient for centromere function. They are however consistent with the idea that tandem repeat organization is a prerequisite for the development of heterochromatin, and once formed, centromeric function is transmitted epigenetically (Houben et al. 2007).

23.1.3.2 Telomeric Region

The termini of barley, rye, and wheat chromosomes were shown to be capped by an array of telomeric repeats strongly hybridizing with the *Arabidopsis thaliana* TTTAGGG telomeric sequence (Schwarzacher and Heslop-Harrison 1991; Werner et al. 1992; Roder et al. 1993; Cheung et al. 1994). Arrays of wheat telomeric sequences were cloned and sequenced and shown to contain the anticipated TTTAGGG sequence along with its degenerate variants (Mao et al. 1997). The arrays of telomeric sequences are more than 25 kb in barley (Roder et al. 1993), between 8 and 50 kb in rye, and between 15 and 175 kb in wheat (Vershinin and Heslop-Harrison 1998). Telomeric sequences were shown to be added *de novo* to broken wheat chromosome ends, presumably during gametogenesis (Werner et al. 1992). Arrays of telomeric sequences are occasionally located interstitially (Cheung et al. 1994).

In *Secale*, *Dasypyrum*, and wheat, telomeric chromatin is more rapidly digested with micrococcal nuclease than is the bulk of chromatin and the resulting nucleosomes are 160 bp compared to 175–185 bp in the bulk of chromatin (Vershinin and Heslop-Harrison 1998). Similar observations were made in other plants (Fajkus et al. 2005), which suggests that the chromatin organization of telomeric sequences differs from that of the rest of the chromatin (Vershinin and Heslop-Harrison 1998). The telomere-specific chromatin organization emanates into telomere associated sequences (TAS) which are juxtaposed to the arrays of telomeric repeats in the Triticeae chromosomes and some of these arrays even have the same chromatin organization as the telomeric chromatin (Vershinin and Heslop-Harrison 1998).

The juxtaposition of telomeric repeats and TAS (Kilian et al. 1999; Roder et al. 1993) is probably universal in the Triticeae genomes because tandem TAS were detected in subtelomeric regions in every Triticeae species investigated to date (Apples et al. 1978, 1981, 1989; Bedbrook et al. 1980; Jones and Flavell 1982a; Xin and Appels 1988; Brandes et al. 1995; Castilho and Heslop-Harrison 1995; Vershinin et al. 1995; Prieto et al. 2004; Salina et al. 2006). TAS fall into several families (Fig. 23.1), which are subjected to the divergence, amplification, and deletion process in different lineages of Triticeae, so that different variants of these families and different quantities of them are present in different lineages (Bedbrook et al. 1980; Appels et al. 1989).

TAS were first cloned from *S. cereale* genomic DNA (Bedbrook et al. 1980). Four families defined by repetition periodicities of 120, 480, 610, and 630 bp

and accounting for 8–12% of the total *S. cereale* DNA were isolated (Bedbrook et al. 1980). The 480-bp periodicity family was renamed as the 350-family because of the 350-bp periodicity of these sequences in most Triticeae genomes (Appels and Moran 1984). The 120-bp family, represented by clone *pSc119*, is present in 1.5×10^6 copies in the *S. cereale* genome (Jones and Flavell 1982a). The sequence has been detected in smaller amounts in the subtelomeric regions of wheat, *Aegilops*, *Hordeum*, and *Critesion* chromosomes (Bedbrook et al. 1980; Castilho and Heslop-Harrison 1995; Taketa et al. 2000) as well as in grasses outside of the tribe (Katsiotis et al. 1997). The 350-family is present in 2×10^5 to 1×10^6 copies in the rye genome, depending on the estimate (Appels et al. 1981; Bedbrook et al. 1980) and has also been detected in other Triticeae genera, such as *Agropyron*, *Critesion*, *Australopyrum*, *Pseudoroegneria*, *Psathyrostachys*, and *Leymus* (Xin and Appels 1988; Appels et al. 1989). Arrays of the 120-bp family are often encountered in interstitial sites (Jones and Flavell 1982b; Vershinin et al. 1995) but the remaining three rye TAS families are almost exclusively associated with telomeres (Jones and Flavell 1982b).

Other TAS families have been reported in Triticeae. Sequence *HvT01*, having repeat periodicity of 118-bp and belonging to the *HvRT* family (Belostotsky and Ananiev 1990) is in subtelomeric regions of 13 of the 14 barley telomeres (Roder et al. 1993) and collocated with the *pAS1/pTA1/Afa* family of repeats with a periodicity around 340 bp (Brandes et al. 1995). The *pAS1/pTA1/Afa* TAS family was first cloned from *Ae. tauschii* (Rayburn and Gill 1986) and has since been detected in a number of *Aegilops*, *Triticum*, *Psathyrostachys*, *Critesion*, *Secale*, *Elymus*, and *Leymus* species in addition to *Hordeum* (Metzlaff et al. 1986; Rayburn and Gill 1986; Anamthawat-Jonsson et al. 1990; Nagaki et al. 1995, 1998a,b, 1999). *pAS1/pTA1/Afa* sequences are also at interstitial locations (Rayburn and Gill 1987; Nagaki et al. 1998a, 1999). A TAS family *pLrTai-I* with a periodicity 570-bp was isolated from *Leymus racemosus* (Kishii et al. 1999) and was shown to be also present in *Triticum* and *Aegilops* where it is however present in pericentromeric heterochromatin (Kishii and Tsujimoto 2002). TAS family *pAesKB52* is highly amplified in the *Ae. speltoidea* genome and accounts for about 1% of the total DNA. While this sequence tends to be at interstitial sites in *Ae. speltoidea*, it is in subtelomeric regions in the closely related *Ae. sharonensis* and *Triticum timopheevii* (Anamthawat-Jonsson and Heslop-Harrison 1993; Salina et al. 2006).

The G genome of *T. timopheevii* and the B genome of *Triticum turgidum* were contributed by *Ae. speltoidea* and a species in *Ae. speltoidea* ancestry, respectively (Dvořák and Zhang 1990; Kilian et al. 2007). Both genomes are mostly devoid of the large subterminal heterochromatic blocks characteristic of *Ae. speltoidea* and other species in the section *Sitopsis* and show greatly reduced numbers of copies of the *Spelt1* family located in them (Pestsova et al. 1998). The significance of the diminution of these sequences during evolution of polyploid wheat is unknown.

TAS families are arranged in blocks juxtaposed to each other in subtelomeric regions. In the boundaries between the blocks, repeats are interspersed among each other with spacers between (Alkhimova et al. 2004). It was suggested that nonhomologous end joining (NEJ) between different repeats facilitated by KU70/KU80 proteins and telomeric proteins is responsible for this repeat organization (Alkhimova et al. 2004).

The functions of telomeric repeats to protect the ends of chromosomes is well established (Blackburn 1986; Harper et al. 2004; Scherthan 2007), but the elucidation of the function of TAS sequences has been frustratingly difficult. TAS are ubiquitous in the Triticeae genomes suggesting that they are important. However, the number of copies of a specific repeat family and the spectrum of families greatly vary even among closely related genomes, as documented earlier, leading to an impression that the primary sequence may not be that important. As in the centromeric regions, it may be that it is the tandem repeated structure of TAS that matters and is a prerequisite for heterochromatin development. Suppression of recombination in heterochromatin may stabilize tandem repeats in the telomeric and subtelomeric regions (Grewal and Klar 1997; Grewal and Moazed 2003). Heterochromatin appears to be essential for meiotic chromosome pairing. At the leptotene stage, telomeric repeats bind to the inner membrane of the nuclear envelope with the assistance of meiosis-specific cohesin and other proteins. Trans-membrane fibrils tether the attachment plate to the cytoskeleton and facilitate the movement of chromosomes via cytoplasmic actin (Conrad et al. 2007). The result of this movement is the chromosome bouquet at the leptone-ma-zygonema transition (reviewed by Harper et al. 2004 and Scherthan 2007). Chromosome movement and bouquet formation is essential for chromosome pairing and crossing over. Studies in yeast and mouse established that telomere associated heterochromatin facilitates chromosome movement and bouquet formation and hence may be essential for homologous pairing and recombination (Scherthan 2007). Additionally, telomeric heterochromatin was shown to affect sister chromatid segregation in mitosis (reviewed by Pandita et al. 2007).

23.1.3.3 Interstitial Sites

In wheat, the polypurine/polypyrimidine GA(A/G) satellite is present in small amounts on most of the A- and D-genome chromosomes, but it is absent from the chromosomes of the diploid donors of these genomes (Peacock et al. 1981; Pedersen et al. 1996). Pedersen et al. (1996) suggested that translocation of the polypurine/polypyrimidine satellite sequences from the B genome chromosomes to the A- and D-genome chromosomes took place in polyploid wheat. Ectopic recombination between isodirectional repeats produces a circular intermediate (Fig. 23.2) (Cohen et al. 2003). Although the intermediate is probably usually lost, it may occasionally be inserted at a different site and result in the translocation of satellite sequences onto a nonhomologous chromosome. Non-homologous end joining may be another mechanism of translocation of repeated

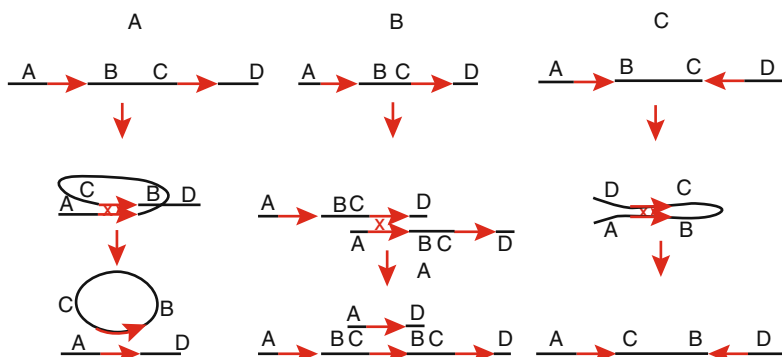


Fig. 23.2 Examples of the fate of genes (B and C) flanked by two repeats as a consequence of ectopic recombination between them. (A) If the repeats are in the isodirectional orientation in the chromosome, intrachromosomal ectopic recombination between them results in the excision of one recombined repeat and intervening DNA between repeats including the B and C genes. (B) If the repeats are in the isodirectional orientation and unequal interchromosomal recombination takes place between sisters or homologues, ectopic recombination results in deletion of the B and C genes in one chromosome and duplication in another. (C) If the repeats are in an inverted orientation, intrachromosomal ectopic recombination between them results in inversion of DNA between the repeats, including the B and C genes (*See Color Insert*)

arrays between nonhomologous chromosomes. Similar processes may also be responsible for exchanges between centromeric and telomeric sequences and the formation of interstitial sites containing telomeric, TAS, and centromeric tandem repeats (Fig. 23.1). The model in Fig. 23.1 implies that intercalated repeats are facultative and derived from the obligatory locations of tandem repeats in the centromeric and telomeric regions. In contrast to the obligatory locations, interstitial sites probably have limited or no biological function.

Several observations support the derived status of interstitial satellite DNA sites. The distances from the centromere of interstitial sites containing polypurine/polypyrimidine repeats are similar in nonhomologous barley chromosomes suggesting their origin via translocation between nonhomologous chromosome arms (Pedersen et al. 1996). It was also observed that within a chromosome, the distance to the centromere of an interstitial site containing TAS in the long arm is often equivalent to the distance from the telomere to the centromere in the short arm, suggesting that the interstitial site in the long arm originated by translocation of tandem array from the telomeric region of the short arm (Fig. 23.1) (Schweizer and Loidl 1987).

23.1.3.4 rRNA Genes

Triticeae nucleolus organizing loci (*Nor* loci) have some properties that are shared with satellite DNA and study of their evolution may be instructive for the understanding of evolution of satellite DNA. The genomes of diploid Triticeae species possess from one to three major *Nor* loci containing hundreds

to thousands of the 18S-5.8S-26S rRNA gene units (rDNA) arranged in tandem (Appels and Honeycutt 1986). In addition, several minor loci can be present per genome (Mukai et al. 1991; Leitch et al. 1992; Jiang and Gill 1994; Dubcovsky and Dvořák 1995). Expressed rDNA units form secondary constrictions. Those that are not expressed form heterochromatin (Mukai et al. 1991; Leitch and Heslop-Harrison 1992). Like satellite DNA, rDNA evolves in concert and loci on different chromosomes share the same sequence variants (Appels and Dvořák 1982; Dvořák and Appels 1982).

Major *Nor* loci are located in homoeologous chromosome groups 1, 5, and 6 across the tribe (summarized by Dubcovsky and Dvořák 1995). A single major rDNA locus per genome is present in rye and *Ae. tauschii* on chromosomes 1 and 5, respectively (Appels et al. 1980; Lawrence and Appels 1986; Lassner et al. 1987). When two loci are present per genome, all possible pairwise combinations have been recorded; they are on chromosomes 1 and 5 in the *T. monococcum*, *T. urartu*, and *Ae. umbellulata* genomes (Gerlach et al. 1980; Miller et al. 1983), on chromosomes 1 and 6 in the *Ae. speltoides* genome and wheat B genome (Crosby 1957; Longwell and Svihla 1960; Dvořák et al. 1984a), and on chromosomes 5 and 6 in the *Lophopyrum elongatum* and barley genomes (Dvořák et al. 1984a; Saghai-Marooft et al. 1984). Three major *Nor* loci per genome were reported in the *Ae. longissima* genome on chromosomes 1, 5, and 6 (Friebe et al. 1993), and the *Psathyrostachys fragilis* genome, in which chromosomes harboring the loci were not identified (Linde-Laursen and Baden 1994).

The location of *Nor* loci on homoeologous chromosome arms in different genomes does not necessarily mean orthology of the loci. For instance, the *Nor* locus on the short arm of *T. monococcum* chromosome 1A^m is nearly 50 cM more distal than the *Nor* locus on the short arm of chromosome 1B, although the two chromosome arms are otherwise collinear (Dubcovsky and Dvořák 1995). Similar changes in the location without structural changes of the arms were demonstrated in other genomes (Dubcovsky and Dvořák 1995).

The following scenario is suggested that accounts for this unusual mobility of Triticeae *Nor* loci and is consistent with the model accounting for the origin of interstitial satellite DNA loci (Schweizer and Loidl 1987). The evolution of a new locus is initiated by translocation of an rDNA repeated array into another location. Homologous and paralogous *Nor* loci are often in physical proximity due to the formation of the nucleolus. rDNA units are therefore usually translocated into another rDNA haplotype, generating a compound haplotype. This compound structure is exemplified by the *Nor-B1* and *Nor-B2* loci present in Chinese Spring wheat (Dvořák and Appels 1986). Sometimes, however, the new site is physically separated from an existing *Nor* locus. Unequal crossing over, random drift, and selection for an increased number of tandem units may result in the evolution of a new major locus and the loss of the original one.

Triticeae 5S*Dna* loci harboring tandem arrays of the 5S rRNA genes evolve like rDNA (Dubcovsky et al. 1996). In *Triticum*, *Aegilops*, and *Lophopyrum*, 5S*Dna* loci are on the short arms of the chromosomes of homoeologous groups 1 and 5 (Dvořák et al. 1989). However, they are on the long arms of

chromosomes 2H and 3H in barley (Kanazin et al. 1993), although chromosomes 2 and 3 are homosequential in the two lineages (Dubcovsky et al. 1996). Like rDNA, the *5SDna* loci translocate into new sites without perturbing collinearity of the chromosomes and both mimic in this attribute satellite DNA.

23.1.4 Interspersed Repeated Sequences

Interspersed repeated sequences are principally transposable elements (TEs) in Triticeae genomes. TEs represent a major component of Triticeae genomes and account for most of the intergenic space. In general terms, TEs fall into two basic classes, retrotransposons (class I) and DNA transposons (class II) (Finnegan 1985). Class I TEs transpose via RNA intermediates (copy and paste mechanism), whereas class II TEs transpose by DNA excision and insertion (cut and paste mechanism). Class I TEs are further subdivided into five subclasses and class II TEs into two subclasses on the basis of their structure and phylogenetic relationships (Wicker et al. 2007). A list of the various TEs found in Triticeae genomes can be found in the Triticeae Repeat Sequence Database (http://wheat.pw.usda.gov/ITMI/Repeats/nrTREP_list.html).

In 3.8 Mb of concatenated BAC DNA studied by Sabot et al. (2005), known TEs represented 54.7% of the total DNA length. However, in a more random sample based on BAC-end sequences of wheat chromosome 3B, known TE families represented 76.3% of total sequenced DNA length (Paux et al. 2006) and in a random plasmid library of *Ae. tauschii*, known and unknown TE families represented 83.4% of the total sequenced DNA length (Li et al. 2004).

Class I TEs, mostly LTR retroelements, were shown to be the most abundant TEs in the *Aegilops* and wheat genomes. They represented 51.2% of the total *Ae. tauschii* genomic DNA (Li et al. 2004) and 67.4% of chromosome 3B DNA (Paux et al. 2006). The most common Class II TEs was the CACTA superfamily, which accounted for 12.3% of the *Ae. tauschii* genome (Li et al. 2004) and 4.9% of chromosome 3B (Paux et al. 2006). Miniature inverted transposable elements (MITEs) accounted for 1% of the *Ae. tauschii* genome (Li et al. 2004) and 0.5% of chromosome 3B (Paux et al. 2006).

Many TEs are nonrandomly distributed along chromosomes. As in maize (SanMiguel et al. 1996), intergenic space is enriched for LTR retroelements showing a complex nested insertion pattern (Wicker et al. 2001, 2003b, 2005). CACTA transposons tend to be associated with genes (Wicker et al. 2003a). MITEs are usually inserted into genes or their immediate vicinity in Triticeae species (Sabot et al. 2005) as in other grasses (Bureau and Wessler 1994). Insertion preferences of TEs undoubtedly account for some of the nonrandomness in the distribution of specific TE families along chromosomes. For example, proximal regions of *Ae. tauschii* chromosomes were shown to be enriched for LTR retrotransposons *Laura* and *Erica* whereas distal regions were enriched for the CACTA family *Caspar* (Li et al. 2004). These patterns are

consistent with the insertion preference of TEs. LTR retrotransposons tend to insert into other LTR retrotransposons in the intergenic regions. The proximal regions of wheat chromosomes tend to be generally gene poor and distal regions generally gene rich (Gill et al. 1993; Akhunov et al. 2003b). Hence, proximal regions are therefore expected to be enriched for LTR retrotransposons and distal regions for class II elements. However, other factors probably modify this basic pattern. *Laura* elements form discrete FISH bands in the proximal regions of *Ae. tauschii* chromosomes (Li et al. 2004), which are difficult to explain by low gene density in the proximal regions alone.

TEs often tend to transpose into their own vicinity (Van Schaik and Brink 1959; Greenblatt and Brink 1962), which gives each chromosome its individual repeated sequence flavor. The distribution of *Lophopyrum* genome-specific interspersed sequence *pLe2* (Zhang and Dvořák 1990) illustrates this process. Clone *pLe2* hybridizes negligibly with Southern blots of wheat DNA. Hybridization of the *pLe2* clone with Southern blots of *L. elongatum* disomic and ditelosomic additions in wheat and *L. ponticum*-wheat recombination substitution lines provide an idea about the distribution and differentiation of the *pLe2* family across *Lophopyrum* genomes (Zhang and Dvořák 1990). The *L. elongatum* genome contains 4×10^6 *pLe2* copies. The sequence is on each chromosome and each chromosome arm in approximately equal amounts. Some variants are distributed across the entire genome whereas others are limited to a single chromosome or chromosome arm. Within a chromosome arm, some subfamilies are spread across the entire arm whereas others are clustered, indicating that they tend to transpose within their own vicinity and give regions within an arm their unique flavors (Zhang and Dvořák 1990).

23.2 Genome Evolution

23.2.1 TEs and Triticeae Genome Evolution

Because LTR retroelements transpose by the “copy and paste” mechanism, the number of sequences within the Class I TE families is expected to continuously increase in a genome (Bennetzen and Kellogg 1997). This tendency is counteracted by sequence deletions. Ectopic LTR recombination results in the excision of the element leaving a solo LTR in place (Vicient et al. 1999; Devos et al. 2002). Ectopic recombination can also take place between neighboring TEs of the same type. Various outcomes of such ectopic recombination are possible and several are illustrated in Fig. 23.2. If duplicated sequences are in the same orientation, the result can be a deletion of all sequences between them on a circular molecule (2A), which is either lost or re-inserted elsewhere, or gene deletion in one chromosome and duplication in another (2B). If they are in an inverted orientation, the result may be an inversion of gene order between the sequences (2C).

A balance between new TE insertions and TE deletions is the principal determinant of the amount of interspersed repeated DNA in a genome. In turn, the balance is also the principal determinant of genome size, because interspersed DNA represents the largest component of Triticeae genomes. The genome will grow if the balance is in favor of insertions and will contract if it is in favor of deletions. As illustrated with the distribution of *pLe2* family across the *L. elongatum* genome, the acquisition of interspersed repeated DNA is quantitatively similar among individual chromosome arms. Consequently, the arm ratio and relative sizes of the seven chromosomes in Triticeae genomes remain nearly constant even when they differ considerably in the overall genome size (Dvořák et al. 1984b). Only chromosome arms that were involved in asymmetric translocations and inversions deviate from the consensus karyotype (Dvořák et al. 1984b).

23.2.2 Gene Order Paradox

Not only is the basic morphology of Triticeae genomes remarkably conservative, so is the gene order. Comparative mapping of the barley and *T. monococcum* genomes suggested that the two genomes differ by two large paracentric inversions, four putative inversions of two markers each, and one translocation (Dubcovsky et al. 1996). Although it would be surprising if additional gross changes in gene order between wheat and barley genomes would not be discovered if denser comparative maps were available, the low number of such differences discovered is nevertheless surprising given the fact that the two genomes diverged about 10 million years ago (Huang et al. 2002; Ramakrishna et al. 2002; Dvořák and Akhunov 2005). The conservation of gene order between the wheat and barley genomes is remarkable when viewed against the background of the turnover rate of intergenic DNA. Intergenic DNA is almost entirely replaced by deletions and insertions in about 3 million years in Triticeae genomes (Dubcovsky and Dvořák 2007). Hence, during the 10 million years since the divergence of wheat and barley, most of the intergenic DNA was replaced several times in each lineage. The high rate of intergenic DNA turnover is a characteristic feature of large Triticeae genomes, and it is nearly two orders of magnitude greater than the turnover rate in mammalian genomes (Dubcovsky and Dvořák 2007).

Given this high rate of sequence turnover in the intergenic regions, gene excisions, inversions, and translocations caused by ectopic recombination (Fig. 23.2) should obliterate gene synteny between Triticeae genomes. Instead, the order of genes is conserved, not only across the many Triticeae genomes, but, to a large extent also across the entire grass family (Gale and Devos 1998; Sorrells et al. 2003). The contrast between the low level of synteny expected on the basis of rapid turnover of intergenic DNA and the level of synteny actually observed in Triticeae genomes is a paradox.

23.2.3 *Conservative and Dynamic Strata of Triticeae Genomes*

To account for this paradox, it is suggested that Triticeae genomes consist of two strata, “conservative” and “dynamic,” greatly differing in evolutionary stability. The conservative stratum consists principally of genes and surrounding single-copy DNA. The dynamic stratum consists of TEs forming the intergenic DNA and duplicated genes and gene fragments recently embedded into it. It is assumed that ectopic recombination between repeats is the most frequent cause of gross structural changes in chromosomes (Fig. 23.2). The likelihood of ectopic recombination decreases with distance between repeats (Lassner and Dvořák 1986), and ectopic recombination therefore preferentially involves TEs that are near each other. Most deletions are consequently short, i.e., less than the average distance between genes in Triticeae genomes (Wicker et al. 2003b; Dvořák et al. 2006; Dubcovsky and Dvořák 2007). Most deletions will therefore delete only TEs. Deletions taking place in a gene cluster may however involve from one to several genes (Chantret et al. 2005). Deletions in the intergenic regions are largely unconstrained by purifying selection allowing that space to turn over with the high rate reported by Dubcovsky and Dvořák (2007). However, deletions of unique genes are largely deleterious in diploids and purifying selection acts against them (Dvořák and Akhunov 2005), preserving *status quo*, i.e., synteny. It is suggested that variation in the strength of purifying selection acting on deletions in the dynamic and conservative strata is responsible for their contrasting stability.

The contrasting stability of genes in the conservative and dynamic strata is illustrated the aci-reductone dioxygenase (*ALT*) gene family. This family evolved within the past two million years in the einkorn wheat lineage (Akhunov et al. 2007a). The ancestral gene of the family (*ALT-1*) is a single-copy gene located on chromosome 1A and its location is conserved across grasses (Akhunov et al. 2007a). In diploid species of Triticeae, single-copy genes are duplicated at a rate of $2.9 \times 10^{-3} \text{ locus}^{-1} \text{ MY}^{-1}$ (Dvořák and Akhunov 2005). The duplication rate of *ALT-1* is consistent with this rate. With two exceptions, the einkorn lineage being one of them, *ALT-1* remained as a single-copy gene across the *Aegilops-Triticum* alliance and across the Triticeae (Akhunov et al. 2007a). *ALT-1* is stable, its synteny is conserved across the grass family, and it behaves in all respects as a gene in the conservative stratum of Triticeae genomes. Duplication of *ALT-1* in the einkorn lineage generated a duplicated gene on chromosome 4A. The duplicated gene duplicated sequentially three times; always it was the most recently produced copy that duplicated again. The duplication rate of these genes was nearly three orders of magnitude greater than that of *ALT-1* (Akhunov et al. 2007a). This accelerated duplication rate was caused by insertions of the duplicated genes into two different TEs in the intergenic DNA (Akhunov et al. 2007a). The duplicated *ALT* genes became dispersed across the einkorn genome, are out of synteny, and behaved as representative of the dynamic stratum.

23.2.4 *Recombination and Gene Content Evolution Along the Centromere-Telomere Axis of Triticeae Chromosomes*

The frequency of meiotic recombination increases along the centromere-telomere axis of chromosome arms in all Triticeae genomes investigated to date (Dvořák and Chen 1984; Gill et al. 1993; Lukaszewski and Curtis 1993; Kunzel et al. 2000; Akhunov et al. 2003b). In the *Ae. tauschii* genome, regions of low- and high-recombination rates alternate along most chromosome arms and are superimposed on this basic recombination rate gradient (J. Dvořák, unpublished). The causes of the recombination rate gradient and minima and maxima superimposed on it are poorly understood. Centromeres are known to suppress recombination in their vicinity (Thompson 1964). It seems unlikely that the centromere effect is the only, or even the primary, mechanism, because a great deal of variation exists among closely related Triticeae species in recombination rate distribution. For example, the genomes of *Ae. speltooides* and *T. monococcum* are of similar sizes and have chromosomes with similar morphology. However, the recombination rate gradient along the centromere-telomere axis is much steeper in the *Ae. speltooides* chromosomes than in the *T. monococcum* chromosomes (Luo et al. 2005). Multiple meiotic crossovers are rare in *Ae. speltooides* and proximal regions that are nearly 100 cM long in *T. monococcum* are contracted to only several cM in *Ae. speltooides* (Luo et al. 2005). Distal regions that are only several cM long in *T. monococcum*, are nearly 50 cM in a number of *Ae. speltooides* chromosome arms.

A number of aspects of gene content correlate with recombination rate and the position of genes along the centromere-telomere axis. Single-copy genes are preferentially located in the proximal, low-recombination regions, whereas multigene loci are preferentially located in distal, high-recombination regions (Akhunov et al. 2003b). Gene density along the centromere-telomere axis also correlates with recombination rate (Akhunov et al. 2003b; Dvořák et al. 2003). This relationship will be explored in greater detail below. Restriction fragment length polymorphism (RFLP) and single nucleotide polymorphism (SNP) in diploid *Aegilops* species and polyploid wheat increase along the centromere-telomere axis and this increase correlates with recombination rates (Dvořák et al. 1998; Akhunov et al. 2007b). Deletions of genes and associated single-copy DNA in tetraploid and hexaploid wheat were found to be almost exclusively located in distal, high-recombination regions (Dvořák et al. 2004). Recently duplicated genes tend to be preferentially located in distal, high recombination regions and these regions appear therefore as the cradles of genetic novelty in wheat genomes (Dvořák and Akhunov 2005). Wheat genes not detected by homology search in the rice genome are also most often located in distal regions of wheat chromosomes (See et al. 2006). Synteny among the three wheat genomes and between the wheat genomes and the rice genome is being eroded faster in distal, high-recombination regions and conserved in proximal, low-recombination regions (Akhunov et al. 2003a,b).

23.2.4.1 Variation in Gene Density Along Chromosomes

It was pointed out above that gene density correlates with recombination rate in wheat chromosomes. Proximal, low-recombination regions have low gene density and distal, high-recombination regions have high-gene density (Gill et al. 1993; Akhunov et al. 2003b). A gradient reported by Akhunov et al. (2003b) is an average across all chromosome arms of hexaploid wheat, and as such, it fails to reflect local variation that may exist in specific chromosome arms. Sandhu and Gill (2002) concluded from deletion bin mapping of wheat homoeologous group 1 chromosomes that when examined in single chromosomes, gene-rich and gene-poor regions are scattered across the entire chromosome and gene density in gene-rich regions was estimated to be about one gene per 20 kb. They further suggested that the difference in the overall sizes between Triticeae genomes and the rice genome were mainly caused by the expansion of gene-poor regions in Triticeae genomes by TE insertions. Devos et al. (2005) expressed doubts that estimates of one gene per 5–20 kb are representative of gene-containing space in Triticeae genomes and suggested that one gene per 75 kb is a more realistic estimate.

Sandhu et al. (2001) inferred that the short arm of chromosome 1 contains two gene-rich islands that contained 86% of genes whereas the rest of the arm contained only 14% of genes. The two islands represented 15% of the arm length. If most genes in Triticeae genomes were clustered into several gene-rich islands per arm, most BAC contigs would fall into gene-rich and gene-poor classes. Such bimodality was not observed in contigs assembled from nearly 200,000 *Ae. tauschii* BAC clones and hybridized with nearly 2,000 wheat EST unigenes (Luo et al. 2003). Moreover, the precipitous difference in numbers of genes between the gene-rich and gene-poor regions in the short arm of wheat homoeologous group 1 was not substantiated by genetic mapping. On the genetic map of the *Ae. tauschii* chromosome arm 1DS containing 82 loci mapped with random cDNAs (ESTs), the region of the map corresponding to bins classified as gene-poor by Sandhu et al. (2001) contained 35% of the EST loci (N. Huo, K.R. Deal, M.C. Luo, Y.Q. Gu, J. Dvořák, unpublished) as opposed to 14% reported by Sandhu et al. (2001). The underestimation of genes in the low-gene density region was very likely caused by nonrandom nature of markers. Loci devoid of RFLP are preferentially located in proximal, low-recombination regions of Triticeae chromosomes (Dvořák et al. 1998). Such markers have been excluded from genetic mapping, and the standard set of Triticeae RFLP markers is therefore depleted in proximal, low-recombination regions of Triticeae chromosomes. The genetic and deletion bin maps that have been used as supporting evidence for the existence of gene-rich islands were either not of sufficient density or they suffered from other limitations, as pointed out above for the short arm of Triticeae chromosome 1, which precluded the provision of convincing evidence that most genes are clustered in several gene-rich islands per chromosome arm.

23.2.4.2 The Cause of Correlation Between Gene Density and Recombination Rate

Although the existence of gene-rich islands is questioned, gene density is nevertheless known to increase along the centromere-telomere axis of Triticeae chromosomes and to correlate with recombination rates along wheat chromosome arms. Recombination rates correlate also with deletion rates in wheat (Dvořák et al. 2004; Dvořák and Akhunov 2005). It is suggested here that the latter relationship is responsible for variation in gene density along Triticeae chromosomes. Consider the consequences of deletions taking place in the dynamic stratum of a genome as opposed to those in the conservative stratum. Deletions of repeated sequences are largely selectively neutral while those involving genes are usually under purifying selection and selected against. Without insertions of new TEs, intergenic regions (the dynamic stratum) are therefore destined to contract due to the difference in purifying selection acting on TEs and genes (Fig. 23.3). Since deletions are more frequent in distal, high-recombination regions than in proximal, low-recombination regions of wheat chromosomes (Dvořák et al. 2004; Dvořák and Akhunov 2005), the contraction of intergenic regions will be faster in distal, high-recombination regions than in proximal, low-recombination regions (Fig. 23.3). With time, more intergenic DNA will be deleted from distal, high-recombination regions than from proximal, low-recombination regions and gene density will increase disproportionately more in distal, high-recombination regions than in proximal, low-recombination regions of Triticeae chromosomes and will correlate with recombination rate.

It is predicted by this model that if the recombination rate gradient were reversed, being high near the centromere and low near the telomere, gene density gradient would also reverse. *Allium* shows such a reversal of recombination gradient along the centromere-telomere axis (Khrustaleva et al. 2005). It

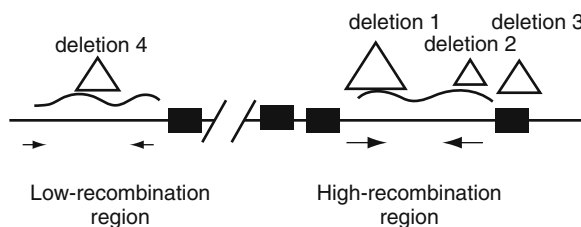


Fig. 23.3 Deletion/selection model accounting for the relationship between gene density and recombination rate. Short horizontal bars represent TEs and the *black boxes* represent genes. The triangles represent deletions; deletion 1, 2, and 4 involve TEs in the intergenic space and deletion 3 involves a gene. The rate and direction of contraction of intergenic space are indicated by sizes and directions of *horizontal arrows*. It is hypothesized that deletions 1, 2, and 4 are neutral whereas deletion 3 is selected against

also shows a reversal of the gene density gradient (Khrustaleva et al. 2005) as predicted by the model proposed here.

23.2.5 *The Evolutionary Significance of Repeated DNA*

A traditional view of repeated nucleotide sequences, particularly those of the interspersed type, is embodied in the parasitic “selfish DNA” hypothesis (Orgel and Crick 1980) or “junk DNA” (Ohno 1980). Evidence is beginning to surface suggesting an adaptive value of TEs (reviewed by Morgante 2006).

Adaptive value of repeated DNA is implicit in the nucleotype theory (Bennett 1972). In diploid organisms, the nuclear and cellular volume, the duration of DNA replication, the duration of mitosis and cell cycle positively correlate with DNA amount per nucleus (Van’t Hof and Sparrow 1963). The nucleotype theory postulates that DNA amount affects fundamental aspects of development, such as the minimum generation time. The nucleotypic effects of the nucleus are outside of the domain of the coding properties of the nucleus. According to the nucleotypic theory, the large sizes of Triticeae genomes limit the rates with which Triticeae plants develop, making them adapted to the temperate zone, where Triticeae species dominate. There may be other nucleotypic effects. For instance, nuclear volume affects cellular water content and environmental stress tolerance (Dvořák and Fowler 1978).

In addition to nucleotypic effects, interspersed repeated sequences may facilitate adaptation by facilitating evolution of new genes. Rice mutator-like transposable elements (*MULEs*) (Jiang et al. 2004) and maize *helitrons* (Kapitonov and Jurka 2001) propagate gene fragments. Although most of them are pseudogenes in these specific cases (Juretic et al. 2006), other less spectacular gene duplications may actually result in new functional genes, as illustrated by the Triticeae *ALP* family (Akhunov et al. 2007a).

Some of the Class II TEs tend to insert near genes. This is particularly true for the CATCA superfamily of TEs. Their insertions can disrupt the natural regulation of genes. They can also provide promoters for novel gene regulation and thus generate new genes (Akhunov et al. 2007a). Since many Triticeae TEs are expressed (Echenique et al. 2002; Li et al. 2004), the dynamic strata of Triticeae genomes contain endless supply of ready-made promoters furnished by TEs that can drive expression of duplicated coding sequences inserted into them or result in gene silencing via the siRNA pathway.

In polyploids, gene deletions generated by ectopic recombination between TEs may facilitate “diploidization” of polyploids (Dvořák et al. 2004). While gene deletions may have dramatic, usually deleterious effects in diploids, they often have minor or no effects in polyploids and may be tolerated (Dvořák et al. 2004; Dvořák and Akhunov 2005). It was hypothesized that gene deletions result in quantitative variation in polyploids enhancing the plasticity of polyploid genomes (Dubcovsky and Dvořák 2007).

It is not difficult to imagine that even a small adaptive benefit of repeated DNA would be meaningful in species with such vast amounts of repeated DNA as species in the tribe Triticeae.

23.3 Conclusions

Triticeae genomes share two conspicuous attributes, large genome sizes caused by vast amounts of repeated DNA and steep recombination gradients along the centromere-telomere axis. Whether or not these variables are causally related is currently unknown. However, there is ample evidence that gene content of Triticeae genomes is structured along the centromere-telomere axis and recombination rate gradients. Numerous genomic variables correlate with recombination rates and gene position on the centromere-telomere axis of chromosomes. The challenge ahead of Triticeae genomics is to determine causalities in those relationships and elucidate underlying mechanisms. Strides are being made in other plant and animal systems to clarify the function of centromeric satellite sequences. However, the role of TAS remains obscured. Chromosomes of diploid and polyploid Triticeae species have large amounts of these sequences, are eminently suitable for complex chromosome manipulations, and are therefore potentially excellent models for the elucidation of the function of this repeated DNA component. Triticeae are also potentially an excellent system for clarification of the role of the interspersed repeated sequences and the dynamic stratum of large plant genomes for the evolution of new genes and genome evolution in general.

Acknowledgment The author is grateful to Patrick E. McGuire and Karin R. Deal for reading the manuscript and making valuable suggestions.

References

- Akhunov, E.D., Akhunov, A.R., Linkiewicz, A.M., Dubcovsky, J., Hummel, D., Lazo, G., Chao, S., Anderson, O.D., David, J., Qi, L.L., Echaliier, B., Gill, B.S., Miftahudin, Gustafson, J.P., La Rota, M., Sorrells, M.E., Zhang, D., Nguyen, H.T., Kalavacharla, V., Hossain, K., Kianian, S., Peng, J., Lapitan, N.L.V., Wennerlind, E.J., Nduati, V., Anderson, J.A., Sidhu, D., Gill, K., McGuire, P.E., Qualset, C.O. and Dvořák, J. (2003a) Synteny perturbations between wheat homoeologous chromosomes by locus duplications and deletions correlate with recombination rates along chromosome arms. *Proc. Natl. Acad. Sci. USA* 100, 10836–10841.
- Akhunov, E.D., Goodyear, J.A., Geng, S., Qi, L.-L., Echaliier, B., Gill, B.S., Miftahudin, Gustafson, J.P., Lazo, G., Chao, S., Anderson, O.D., Linkiewicz, A.M., Dubcovsky, J., La Rota, M., Sorrells, M.E., Zhang, D., Nguyen, H.T., Kalavacharla, V., K., H., Kianian, S.F., Peng, J., Lapitan, N.L.V., Gonzalez-Hernandez, J.L., Anderson, J.A., Choi, D.-W., Close, T.J., Dilbirli, M., Gill, K.S., Walker-Simmons, M.K., Steber, C., McGuire, P.E., Qualset, C.O. and Dvořák, J. (2003b) The organization and rate of evolution of the wheat

- genomes are correlated with recombination rates along chromosome arms. *Genome Res.* 13, 753–763.
- Akhunov, E.D., Akhunova, A.R. and Dvořák, J. (2005) BAC libraries of *Triticum urartu*, *Aegilops speltoides* and *Ae. tauschii*, the diploid ancestors of polyploid wheat. *Theor. Appl. Genet.* 111, 1617–1622.
- Akhunov, E.D., Akhunova, A.R. and Dvořák, J. (2007a) Mechanisms and rates of birth and death of dispersed duplicated genes during the evolution of a multigene family in diploid and tetraploid wheats. *Mol. Biol. Evol.* 24, 539–550.
- Akhunov, E.D., Akhunova, A.R., Saini, B., Grishina, I., Morrell, P.L., Toleno, D., Clegg, M.T. and Dvořák, J. (2007b) Genetic diversity of diploid ancestors of wheat. *Plant and Animal Genome XV*, San Diego, CA, pp. 168.
- Alkhimova, O.G., Mazurok, N.A., Potapova, T.A., Zakian, S.M., Heslop-Harrison, J.S. and Vershinin, A.V. (2004) Diverse patterns of the tandem repeats organization in rye chromosomes. *Chromosoma* 113, 42–52.
- Amor, D.J., Kalitsis, P., Sumer, H. and Choo, K.H.A. (2004) Building the centromere: from foundation proteins to 3D organization. *Trends in Cell Biol.* 14, 359–368.
- Ananthawat-Jonsson, K., Schwarzacher, T., Leitch, A.R., Bennett, M.D. and Heslop-Harrison, J.S. (1990) Discrimination between closely related species using genomic DNA as a probe. *Theor. Appl. Genet.* 79, 721–728.
- Ananthawat-Jonsson, K. and Heslop-Harrison, J.S. (1993) Isolation and characterization of genome-specific DNA sequences in Triticeae species. *Mol. Gen. Genet.* 240, 151–158.
- Appels, R., Driscoll, C. and Peacock, W.J. (1978) Heterochromatin and highly repeated DNA sequences in rye (*Secale cereale*). *Chromosoma* 70, 67–89.
- Appels, R., Gerlach, W.L., Dennis, E.S., Swift, H. and Peacock, W.J. (1980) Molecular and chromosomal organization of DNA sequences coding for the ribosomal RNAs in cereals. *Chromosoma* 78, 293–311.
- Appels, R., Dennis, E.S., Smyth, D.R. and Peacock, W.J. (1981) Two repeated DNA-sequences from the heterochromatic regions of rye (*Secale cereale*) chromosomes. *Chromosoma* 84, 265–277.
- Appels, R. and Dvořák, J. (1982) The wheat ribosomal DNA spacer: its structure and variation in populations and among species. *Theor. Appl. Genet.* 63, 337–348.
- Appels, R. and Moran, L.B. (1984) Molecular analysis of alien chromatin introduced into wheat. In: P.J. Gustafson (Ed.), *Stadler Genetics Symposia*. University of Missouri, Columbia, MO, pp. 529–557.
- Appels, R. and Honeycutt, R.L. (1986) rDNA: evolution over a billion years. In: S.K. Dutta (Ed.), *DNA Systematics*. CRC Press, Florida, pp. 81–135.
- Appels, R., Reddy, P., McIntyre, C.L., Moran, L.B., Frankel, O.H. and Clarke, B.C. (1989) The molecular-cytogenetic analysis of grasses and its application to studying relationships among species Triticeae. *Genome* 31, 122–133.
- Aragon-Alcaide, L., Miller, T., Schwarzacher, T., Reader, S. and Moore, G. (1996) A cereal centromeric sequence. *Chromosoma* 105, 261–268.
- Arumuganathan, K. and Earle, E.D. (1991) Nuclear DNA content of some important plant species. *Plant Mol. Biol. Rep.* 9, 208–218.
- Bedbrook, J.R., Jones, J., O'Dell, M., Thompson, R.D. and Flavell, R.B. (1980) A molecular description of telomeric heterochromatin in *Secale* species. *Cell* 19, 545–560.
- Belostotsky, D.A. and Ananiev, E.V. (1990) Characterization of relic DNA from barley genome. *Theor. Appl. Genet.* 80, 374–380.
- Bennett, M.D. (1972) Nuclear DNA content and minimum generation time in herbaceous plants. *Proc. Royal Soc. Lond. B* 181, 109–135.
- Bennett, M.D. and Smith, J.B. (1976) Nuclear DNA amounts in angiosperms. *Phil. Trans. Royal Soc. Lond. Ser. B, Biol. Sci.* 274, 227–274.
- Bennett, M.D., Gustafson, J.P. and Smith, J.B. (1977) Variation in nuclear DNA in genus *Secale*. *Chromosoma* 61, 149–176.

- Bennetzen, J.L. and Kellogg, E.A. (1997) Do plants have a one-way ticket to genomic obesity? *Plant Cell* 9, 1509–1514.
- Bernard, P., Maure, J.F., Partridge, J.F., Genier, S., Javerzat, J.P. and Allshire, R.C. (2001) Requirement of heterochromatin for cohesion at centromeres. *Science* 294, 2539–2542.
- Blackburn, E.H. (1986) Structure and formation of telomeres in Holotrichous Ciliates. *Internatl. Rev. Cytol.* 99, 29–47.
- Brandes, A., Roder, M.S. and Ganai, M.W. (1995) Barley telomeres are associated with two different types of satellite DNA-sequences. *Chromosome Res.* 3, 315–320.
- Bureau, T.E. and Wessler, S.R. (1994) Mobile inverted-repeat elements of the *tourist* family are associated with the genes of many cereal grasses. *Proc. Natl. Acad. Sci. USA* 91, 1411–1415.
- Castilho, A. and Heslop-Harrison, J.S. (1995) Physical mapping of 5S and 18S-25S rDNA and repetitive DNA sequences in *Aegilops umbellulata*. *Genome* 38, 91–96.
- Chantret, N., Salse, J., Sabot, F., Rahman, S., Bellec, A., Laubin, B., Dubois, I., Dossat, C., Sourdille, P., Joudrier, P., Gautier, M.F., Cattolico, L., Beckert, M., Aubourg, S., Weissenbach, J., Caboche, M., Bernard, M., Leroy, P. and Chalhou, B. (2005) Molecular basis of evolutionary events that shaped the hardness locus in diploid and polyploid wheat species (*Triticum* and *Aegilops*). *Plant Cell* 17, 1033–1045.
- Cheng, Z.J. and Murata, M. (2003) A centromeric tandem repeat family originating from a part of *Ty3/gypsy* retroelement in wheat and its relatives. *Genetics* 164, 665–672.
- Cheung, W.Y., Money, T.A., Abbo, S., Devos, K.M., Gale, M.D. and Moore, G. (1994) A family of related sequences associated with (TTTAGGG)_n repeats are located in the interstitial regions of wheat chromosomes. *Mol. Gen. Genet.* 245, 349–354.
- Cohen, S., Yacobi, K. and Segal, D. (2003) Extrachromosomal circular DNA of tandemly repeated genomic sequences in *Drosophila*. *Genome Res.* 13, 1133–1145.
- Conrad, M.N., Lee, C.Y., Wilkerson, J.L. and Dresser, M.E. (2007) MPS3 mediates meiotic bouquet formation in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* 104, 8863–8868.
- Crosby, A.R. (1957) Nucleolar activity of lagging chromosomes in wheat. *Amer. J. Bot.* 44, 813–822.
- Cuadrado, A., Schwarzacher, T. and Jouve, N. (2000) Identification of different chromatin classes in wheat using in situ hybridization with simple sequence repeat oligonucleotides. *Theor. Appl. Genet.* 101, 711–717.
- Cuadrado, A. and Jouve, N. (2007) The nonrandom distribution of long clusters of all possible classes of trinucleotide repeats in barley chromosomes. *Chromos. Res.* 15, 711–720.
- Dennis, E.S., Gerlach, W.L. and Peacock, W.J. (1980) Identical polypyrimidine-polypurine satellite DNAs in wheat and barley. *Heredity* 44, 344–366.
- Devos, K.M., Brown, J.K.M. and Bennetzen, J.L. (2002) Genome size reduction through illegitimate recombination counteracts genome expansion in *Arabidopsis*. *Genome Res.* 12, 1075–1079.
- Devos, K.M., Ma, J.X., Pontaroli, A.C., Pratt, L.H. and Bennetzen, J.L. (2005) Analysis and mapping of randomly chosen bacterial artificial chromosome clones from hexaploid bread wheat. *Proc. Natl. Acad. Sci. USA* 102, 19243–19248.
- Dubcovsky, J. and Dvořák, J. (1995) Ribosomal RNA loci: nomads in the Triticeae genomes. *Genetics* 140, 1367–1377.
- Dubcovsky, J., Luo, M.C., Zhong, G.Y., Bransteitter, R., Desai, A., Kilian, A., Kleinhofs, A. and Dvořák, J. (1996) Genetic map of diploid wheat, *Triticum monococcum* L., and its comparison with maps of *Hordeum vulgare* L. *Genetics* 143, 983–999.
- Dubcovsky, J. and Dvořák, J. (2007) Genome plasticity a key factor in the success of polyploid wheat under domestication. *Science* 316, 1862–1866.
- Dvořák, J. and Fowler, D.B. (1978) Cold hardiness potential of triticale and tetraploid rye. *Crop Sci.* 17, 477–478.

- Dvořák, J. and Appels, R. (1982) Chromosomal and nucleotide sequence differentiation in genomes of polyploid *Triticum* species. *Theor. Appl. Genet.* 63, 349–360.
- Dvořák, J. and Chen, K.-C. (1984) Distribution of nonstructural variation between wheat cultivars along chromosome arm *6Bp*: evidence from the linkage map and physical map of the arm. *Genetics* 106, 325–333.
- Dvořák, J., Lassner, M.W., Kota, R.S. and Chen, K.C. (1984a) The distribution of the ribosomal RNA genes in the *Triticum speltoides* and *Elytrigia elongata* genomes. *Can. J. Genet. Cytol.* 26, 628–632.
- Dvořák, J., McGuire, P.E. and Mendlinger, S. (1984b) Inferred chromosome morphology of the ancestral genome of *Triticum*. *Plant Syst. Evol.* 144, 209–220.
- Dvořák, J. and Appels, R. (1986) Investigation of homologous crossing over and sister chromatid exchange in the wheat *Nor-2* locus coding for rRNA and *Gli-B2* locus coding for gliadins. *Genetics* 113, 1037–1056.
- Dvořák, J., Zhang, H.B., Kota, R.S. and Lassner, M. (1989) Organization and evolution of the 5S ribosomal RNA gene family in wheat and related species. *Genome* 32, 1003–1016.
- Dvořák, J. and Zhang, H.B. (1990) Variation in repeated nucleotide sequences sheds light on the phylogeny of the wheat B and G genomes. *Proc. Natl. Acad. Sci. USA* 87, 9640–9644.
- Dvořák, J., Luo, M.-C. and Yang, Z.-L. (1998) Restriction fragment length polymorphism and divergence in the genomic regions of high and low recombination in self-fertilizing and cross-fertilizing *Aegilops* species. *Genetics* 148, 423–434.
- Dvořák, J., Akhunov, A.D., Akhunov, A.R., Luo, M.C., Linkiewicz, A.M., Dubcovsky, J., Hummel, D., Lazo, G., Chao, S., Anderson, O.D., David, J., Qi, L.L., Echalié, B., Gill, B.S., Miftahudin, Gustafson, J.P., La Rota, M., Sorrells, M.E., Zhang, D., Nguyen, H.T., Kalavacharla, V., Hossain, K., Kianian, S.F., Peng, J.H., Lapitan, N.L.V., Wennerlind, E.J., Nduati, V., Anderson, J.A., Sidhu, D., Gill, K.S., Choi, D.-W., Close, T.J., McGuire, P.E. and Qualset, C.O. (2003) New insights into the organization and evolution of wheat genomes. In: N.E. Pogna, M. Romano, E.A. Pogna and G. Galterio (Eds.), 10th International Wheat Genetics Symposium. SIMI, Rome, Italy, pp. 261–264.
- Dvořák, J., Yang, Z.-L., You, F.M. and Luo, M.C. (2004) Deletion polymorphism in wheat chromosome regions with contrasting recombination rates. *Genetics* 168, 1665–1675.
- Dvořák, J. and Akhunov, E.D. (2005) Tempos of deletions and duplications of gene loci in relation to recombination rate during diploid and polyploid evolution in the *Aegilops-Triticum* alliance. *Genetics* 171, 323–332.
- Dvořák, J., Akhunov, E.D., Akhunov, A.R., Deal, K.R. and Luo, M.C. (2006) Molecular characterization of a diagnostic DNA marker for domesticated tetraploid wheat provides evidence for gene flow from wild tetraploid wheat to hexaploid wheat. *Mol. Biol. Evol.* 23, 1386–1396.
- Echenique, V.C., Stamova, B., Volters, P., Lazo, G.R., Carollo, V.L. and Dubcovsky, J. (2002) Frequencies of *Ty1-copia* and *Ty3-gypsy* retroelements within the *Triticeae* EST databases. *Theor. Appl. Genet.* 104, 840–844.
- Fajkus, J., Sykorova, E. and Leitch, A.R. (2005) Telomeres in evolution and evolution of telomeres. *Chromos. Res.* 13, 469–479.
- Finnegan, D.J. (1985) Transposable elements in eukaryotes. *Internatl. Rev. Cytol.* 93, 281–326.
- Flavell, R.B., Bennett, M.D., Smith, J.B. and Smith, D.B. (1974) Genome size and the proportion of repeated nucleotide sequence DNA in plants. *Biochem. Gen.* 12, 257–269.
- Friebe, B., Tuleen, N., Jiang, J. and Gill, B.S. (1993) Standard karyotype of *Triticum longissimum* and its cytogenetic relationship with *T. aestivum*. *Genome* 36, 731–742.
- Fukui, K.N., Suzuki, G., Lagudah, E.S., Rahman, S., Appels, R., Yamamoto, M. and Mukai, Y. (2001) Physical arrangement of retrotransposon-related repeats in centromeric regions of wheat. *Plant Cell Physiol.* 42, 189–196.
- Gale, M.D. and Devos, K.M. (1998) Comparative genetics in the grasses. *Proc. Natl. Acad. Sci. USA* 95, 1971–1974.

- Gerlach, W.L., Appels, R., Dennis, E.S. and Peacock, W.J. (1979) Evolution and analysis of wheat genomes using highly repeated DNA sequences. In: S. Ramanujam (Ed.), Fifth International Wheat Genetics Symposium. Indian Society of Genetics and Plant Breeding, Indian Agricultural Research Institute, New Delhi, India, pp. 81–91.
- Gerlach, W.L., Miller, T.E. and Flavell, R.B. (1980) The nucleolus organizers of diploid wheats revealed by in situ hybridization. *Theor. Appl. Genet.* 58, 97–100.
- Gill, B.S. and Kimber, G. (1974) Giemsa C-banding and the evolution of wheat. *Proc. Natl. Acad. Sci. USA* 71, 4086–4090.
- Gill, K.S., Gill, B.S. and Endo, T.R. (1993) A chromosome region-specific mapping strategy reveals gene-rich telomeric ends in wheat. *Chromosoma* 102, 374–381.
- Greenblatt, I.M. and Brink, R.A. (1962) Twin mutations in medium variegated pericarp maize. *Genetics* 47, 489–501.
- Grewal, S.I.S. and Klar, A.J.S. (1997) A recombinationally repressed region between *mat2* and *mat3* loci shares homology to centromeric repeats and regulates directionality of mating-type switching in fission yeast. *Genetics* 146, 1221–1238.
- Grewal, S.I.S. and Moazed, D. (2003) Heterochromatin and epigenetic control of gene expression. *Science* 301, 798–802.
- Harper, L., Golubovskaya, I. and Cande, W.Z. (2004) A bouquet of chromosomes. *J. Cell Sci.* 117, 4025–4032.
- Houben, A., Schroeder-Reiter, E., Nagaki, K., Nasuda, S., Wanner, G., Murata, M. and Endo, T.R. (2007) CENH3 interacts with the centromeric retrotransposon cereba and GC-rich satellites and locates to centromeric substructures in barley. *Chromosoma* 116, 275–283.
- Huang, S., Sirikhachornkit, A., Su, X., Faris, J., Gill, B.S., Haselkorn, R. and Gornicki, P. (2002) Genes encoding plastid acetyl-CoA carboxylase and 3-phosphoglycerate kinase of the *Triticum/Aegilops* complex and the evolutionary history of polyploid wheat. *Proc. Natl. Acad. Sci. USA* 99, 8133–8138.
- Hudakova, S., Michalek, W., Presting, G.G., Ten Hoopen, R., Dos Santos, K., Jasencakova, Z. and Schubert, I. (2001) Sequence organization of barley centromeres. *Nucl. Acid Res.* 29, 5029–5035.
- Jakob, S.S., Meister, A. and Blattner, F.R. (2004) The considerable genome size variation of *Hordeum* species (Poaceae) is linked to phylogeny, life form, ecology, and speciation rates. *Mol. Biol. Evol.* 21, 860–869.
- Jiang, J. and Gill, B.S. (1994) New 18S-26S ribosomal RNA gene loci: chromosomal landmarks for the evolution of polyploid wheats. *Chromosoma* 103, 179–185.
- Jiang, N., Bao, Z., Zhang, X., Eddy, S.R. and Wessler, S.R. (2004) Pack-MULE transposable elements mediate gene evolution in plants. *Nature* 30, 569–573.
- Jones, J.D.G. and Flavell, R.B. (1982a) The structure, amount and chromosomal localization of defined repeated DNA-sequences in species of the genus *Secale*. *Chromosoma* 86, 613–641.
- Jones, J.D.G. and Flavell, R.B. (1982b) The mapping of highly-repeated DNA families and their relationship to C-bands in chromosomes of *Secale cereale*. *Chromosoma* 86, 595–612.
- Juretic, N., Hoen, D.R., Huynh, M.L., Marrison, P.M. and Bureau, T.E. (2006) The evolutionary fate of MULE-mediated duplications of host gene fragments in rice. *Genome Res.* 15, 1292–1297.
- Kanazin, V., Ananiev, E. and Blake, T. (1993) The genetics of 5S rRNA encoding multigene families in barley. *Genome* 36, 1023–1028.
- Kapitonov, V.V. and Jurka, J. (2001) Rolling-circle transposons in eukaryotes. *Proc. Natl. Acad. Sci. USA* 98, 8714–8719.
- Katsiotis, A., Hagidimitriou, M. and Heslop-Harrison, J.S. (1997) The close relationship between the A and B genomes in *Avena L.* (Poaceae) determined by molecular cytogenetic analysis of total genomic, tandemly and dispersed repetitive DNA sequences. *Annals Bot.* 79, 103–109.

- Khrustaleva, L.I., de Melo, P.E., van Heusden, A.W. and Kik, C. (2005) The integration of recombination and physical maps in a large-genome monocot using haploid genome analysis in a trihybrid *Allium* population. *Genetics* 169, 1673–1685.
- Kilian, A., Kudrna, D. and Kleinhofs, A. (1999) Genetic and molecular characterization of barley chromosome telomeres. *Genome* 42, 412–419.
- Kilian, B., Özkan, H., Deusch, O., Effgen, S., Brandolini, A., Kohl, J., Martin, W. and Salamini, F. (2007) Independent wheat B and G genome origins in outcrossing *Aegilops* progenitor haplotypes. *Mol. Biol. Evol.* 24, 217–227.
- Kishii, M., Nagaki, K., Tsujimoto, H. and Sasaki, T. (1999) Exclusive localization of tandem repetitive sequences in subtelomeric heterochromatin regions of *Leymus racemosus* (Poaceae, Triticeae). *Chromosome Res.* 7, 519–529.
- Kishii, M., Nagaki, K. and Tsujimoto, H. (2001) A tandem repetitive sequence located in the centromeric region of common wheat (*Triticum aestivum*) chromosomes. *Chromos. Res.* 9, 417–428.
- Kishii, M. and Tsujimoto, H. (2002) Genus-specific localization of the Tail family of tandem-repetitive sequences in either the centromeric or subtelomeric regions in Triticeae species (Poaceae) and its evolution in wheat. *Genome* 45, 946–955.
- Kunzel, G., Korzun, L. and Meister, A. (2000) Cytologically integrated physical restriction fragment length polymorphism maps for the barley genome based on translocation break-points. *Genetics* 154, 397–412.
- Lassner, M. and Dvořák, J. (1986) Preferential homogenization between adjacent and alternate subrepeats in wheat rDNA. *Nucl. Acid. Res.* 14, 5499–5512.
- Lassner, M., Anderson, O. and Dvořák, J. (1987) Hypervariation associated with a 12-nucleotide direct repeat and inferences on intragenomic homogenization of ribosomal RNA gene spacer based on the DNA sequence of a clone from the wheat *Nor-D3* locus. *Genome* 29, 770–781.
- Lawrence, G.J. and Appels, R. (1986) Mapping the nucleolus organizing region, seed protein loci and isozyme loci on chromosome 1R in rye. *Theor. Appl. Genet.* 71, 742–749.
- Leitch, A.R., Mosgoller, W., Shi, M. and Heslop-Harrison, J.S. (1992) Different patterns of rDNA organization at interphase in nuclei of wheat and rye. *J. Cell Sci.* 101, 751–757.
- Leitch, J.J. and Heslop-Harrison, J.S. (1992) Physical mapping of the 18S-5.8S-26S rRNA genes in barley by *in situ* hybridization. *Genome* 35, 1013–1018.
- Li, W., Zhang, P., Fellers, J.P., Friebe, B. and Gill, B.S. (2004) Sequence composition, organization, and evolution of the core Triticeae genome. *Plant J.* 40, 500–511.
- Linde-Laursen, I. and Baden, C. (1994) Comparison of the Giemsa C-banded karyotypes of the three subspecies of *Psathyrostachys fragilis*, subspp. *villosus* (2x), *secaliformis* (2x, 4x), and *fragilis* (2x) (Poaceae), with notes on chromosome pairing. *Pl. Syst. Evol.* 191, 183–198.
- Longwell, A.C. and Svihla, G. (1960) Specific chromosomal control of the nucleolus and of cytoplasm in wheat. *Exp. Cell Res.* 20, 294–312.
- Löve, A. (1984) Conspectus of the Triticeae. *Feddes Repertorium* 95, 425–521.
- Lukaszewski, A.J. and Curtis, C.A. (1993) Physical distribution of recombination in B-genome chromosomes of tetraploid wheat. *Theor. Appl. Genet.* 84, 121–127.
- Luo, M.C., Thomas, C., Deal, K.R., You, F.M., Anderson, O.D., Gu, Y.G., Li, W., Kuruparth, V., Gill, B.S., McGuire, P.E. and Dvořák, J. (2003) Construction of contigs of *Ae. tauschii* genomic DNA fragments cloned in BAC and BiBAC vectors. In: N.E. Pogna, M. Romano, E.A. Pogna and G. Galterio (Eds.) 10th International Wheat Genetics Symposium. Institute Sperimentale per la Cerealicoltura, Roma, Italy, pp. 293–296.
- Luo, M.C., Deal, K.R., Young, Z.L. and Dvořák, J. (2005) Comparative genetic maps reveal extreme crossover localization in the *Aegilops speltoides* chromosomes. *Theor. Appl. Genet.* 111, 1098–1106.
- Mao, L., Devos, K.M., Zhu, L. and Gale, M.D. (1997) Cloning and genetic mapping of wheat telomere-associated sequences. *Mol. Gen. Genet.* 254, 584–591.

- Metzlaff, M., Troebner, W., Baldauf, F., Schlegel, R. and Cullum, J. (1986) Wheat specific repetitive DNA-sequences – construction and characterization of 4 different genomic clones. *Theor. Appl. Genet.* 72, 207–210.
- Miller, J.T., Dong, F.G., Jackson, S.A., Song, J. and Jiang, J.M. (1998) Retrotransposon-related DNA sequences in the centromeres of grass chromosomes. *Genetics* 150, 1615–1623.
- Miller, T.E., Hutchinson, J. and Reader, S.M. (1983) The identification of the nucleolus organiser chromosomes of diploid wheat. *Theor. Appl. Genet.* 65, 145–147.
- Morgante, M. (2006) Plant genome organisation and diversity: the year of the junk! *Curr. Opin. Biotech.* 17, 168–173.
- Mukai, Y., Endo, T.R. and Gill, B.S. (1991) Physical mapping of the 18S.26S rRNA multi-gene family in common wheat: identification of a new locus. *Chromosoma* 100, 71–78.
- Nagaki, K., Tsujimoto, H., Isono, K. and Sasakuma, T. (1995) Molecular characterization of a tandem repeat, Afa family, and distribution among Triticeae. *Genome* 38, 479–486.
- Nagaki, K., Tsujimoto, H. and Sasakuma, T. (1998a) Dynamics of tandem repetitive Afa-family sequences in Triticeae, wheat-related species. *J. Mol. Evol.* 47, 183–189.
- Nagaki, K., Tsujimoto, H. and Sasakuma, Y. (1998b) H genome specific repetitive sequence, pEt2, of *Elymus trachycaulus* in part of Afa family of Triticeae. *Genome* 41, 134–136.
- Nagaki, K., Kishii, M., Tsujimoto, H. and Sasakuma, T. (1999) Tandem repetitive Afa-family sequences from *Leymus racemosus* and *Psathyrostachys juncea* (Poaceae). *Genome* 42, 1258–1260.
- Nasuda, S., Hudakova, S., Schubert, I., Houben, A. and Endo, T.R. (2005) Stable barley chromosomes without centromeric repeats. *Proc. Natl. Acad. Sci. USA* 102, 9842–9847.
- Ohno, S. (1980) So much 'junk' DNA in our genome. *Brookhaven Symp. Biol.* 23, 604–607.
- Orgel, L.E. and Crick, F.H.C. (1980) Selfish DNA – the ultimate parasite. *Nature* 284, 604–607.
- Pandita, T.K., Hunt, C.R., Sharma, G.G. and Yang, Q. (2007) Regulation of telomere movement by telomere chromatin structure. *Cell. Mol. Life Sci.* 64, 131–138.
- Paux, E., Roger, D., Badaeva, E., Gay, G., Bernard, M., Sourdille, P. and Feuillet, C. (2006) Characterizing the composition and evolution of homeologous genomes in hexaploid wheat through BAC-end sequencing on chromosome 3B. *Plant J.* 48, 463–474.
- Peacock, W.J., Gerlach, W.L. and Dennis, E.S. (1981) Molecular aspects of wheat evolution: repeated DNA sequences. In: L.T. Evans and W.J. Peacock (Eds.), *Wheat Science-Today and Tomorrow*. Cambridge University Press, Cambridge, pp. 41–60.
- Pedersen, C., Rasmussen, S.K. and LindeLaursen, I. (1996) Genome and chromosome identification in cultivated barley and related species of the Triticeae (Poaceae) by in situ hybridization with the GAA-satellite sequence. *Genome* 39, 93–104.
- Pestsova, E.G., Goncharov, N.P. and Salina, E.A. (1998) Elimination of a tandem repeat of telomeric heterochromatin during the evolution of wheat. *Theor. Appl. Genet.* 97, 1380–1386.
- Presting, G.G., Malysheva, L., Fuchs, J. and Schubert, I.Z. (1998) A TY3/GYPSY retro-transposon-like sequence localizes to the centromeric regions of cereal chromosomes. *Plant J.* 16, 721–728.
- Prieto, P., Martin, A. and Cabrera, A. (2004) Chromosomal distribution of telomeric and telomeric-associated sequences in *Hordeum chilense* by in situ hybridization. *Hereditas* 141, 122–127.
- Ramakrishna, W., Dubcovsky, J., Park, Y.J., Busso, C., Embereton, J., SanMiguel, P. and Bennetzen, J.L. (2002) Different types and rates of genome evolution detected by comparative sequence analysis of orthologous segments from four cereal genomes. *Genetics* 162, 1389–1400.

- Rayburn, A.L. and Gill, B.S. (1986) Isolation of a D-genome specific repeated DNA sequence from *Aegilops squarrosa*. Plant Mol. Biol. Report. 4, 102–109.
- Rayburn, A.L. and Gill, B.S. (1987) Molecular analysis of the D-genome of the Triticeae. Theor. Appl. Genet. 73, 385–388.
- Roder, M.S., Lapitan, N.L.V., Sorrells, M.E. and Tanksley, S.D. (1993) Genetic and physical mapping of barley telomeres. Mol. Gen. Genet. 238, 294–303.
- Sabot, F., Guyot, R., Wicker, T., Chantret, N., Laubin, B., Chalhoub, B., Leroy, P., Sourdille, P. and Bernard, M. (2005) Updating of transposable element annotations from large wheat genomic sequences reveals diverse activities and gene associations. Mol. Genet. Genom. 274, 119–130.
- Šafář, J., Bartoš, J., Janda, J., Bellec, A., Kubaláková, M., Valárik, M., Pateyron, S., Weiserová, J., Tušková, J., Čihalíková, J., Vrána, J., Šimková, H., Faivre-Rampant, P., Sourdille, P., Caboche, M., Bernard, M., Doležel, J. and Chalhoub, B. (2004) Dissecting large and complex genomes: flow sorting and BAC cloning of individual chromosomes from bread wheat. Plant J. 39, 960–968.
- Saghai-Maroo, M.A., Soliman, K.M., Jorgensen, R.A. and Allard, R.W. (1984) Ribosomal DNA spacer-length polymorphism in barley: Mendelian inheritance, chromosomal location, and population dynamics. Proc. Natl. Acad. Sci. USA 81, 8014–8018.
- Salina, E.A., Lim, K.Y., Badaeva, E.D., Shcherban, A.B., Adonina, I.G., Amosova, A.V., Samatadze, T.E., Vatolina, T.Y., Zoshchuk, S.A. and Leitch, A.R. (2006) Phylogenetic reconstruction of Aegilops section Sitopsis and the evolution of tandem repeats in the diploids and derived wheat polyploids. Genome 49, 1023–1035.
- Sandhu, D., Champoux, J.A., Bondareva, S.N. and Gill, K.S. (2001) Identification and physical localization of useful genes and markers to a major gene-rich region on wheat group 1S chromosomes. Genetics 157, 1735–1747.
- Sandhu, D. and Gill, K.S. (2002) Gene-containing regions of wheat and the other grass genomes. Plant Physiol. 128, 803–811.
- SanMiguel, P., Tikhonov, A., Jin, Y.-K., Moutchoulskaia, N., Zakharov, D., Melake-Berhan, A., Springer, P.S., Edwards, K.J., Lee, M., Avramova, Z. and Bennetzen, J.L. (1996) Nested retrotransposons in the intergenic regions of the maize genome. Science 274, 765–768.
- Scherthan, H. (2007) Telomere attachment and clustering during meiosis. Cellular Mol. Life Sci. 64, 117–124.
- Schwarzacher, T. and Heslop-Harrison, J.S. (1991) *In situ* hybridization to plant telomeres using synthetic oligomers. Genome 34, 317–323.
- Schweizer, D. and Loidl, J. (1987) A model for heterochromatin dispersion and the evolution of C-band patterns. Chromos. Today 9, 61–74.
- See, D.R., Brooks, S., Nelson, J.C., Brown-Guedira, G., Friebe, B. and Gill, B.S. (2006) Gene evolution at the ends of wheat chromosomes. Proc. Natl. Acad. Sci. USA 103, 4162–4167.
- Smith, D.B. and Flavell, R.B. (1975) Characterisation of the wheat genome by renaturation kinetics. Chromosoma 50, 223–242.
- Sorrells, M.E., La Rota, C.M., Bermudez-Kandianis, C.E., Greene, R.A., Kantety, R., Munkvold, J.D., Miftahudin, Mahmoud, A., Gustafson, J.P., Qi, L.L., Echalié, B., Gill, B.S., Matthews, D., Lazo, G., Chao, S., Anderson, O.D., Edwards, H., Linkiewicz, A.M., Dubcovsky, J., Akhunov, E.D., Dvořák, J., Zhang, D., Nguyen, H.T., Peng, J., Lapitan, N.L.V., Gonzalez-Hernandez, J.L., Anderson, J.A., Hossain, K.G., Kalavacharla, V., Kianian, S.F., Choi, D.W., Close, T.J., Dilbirli, M., Gill, K.S., Steber, C., Walker-Simmons, M.K., McGuire, P.E. and Qualset, C.O. (2003) Comparative DNA sequence analysis of wheat and rice genomes. Genome Res. 13, 1818–1827.
- Taketa, S., Ando, H., Takeda, K., Harrison, G.E. and Heslop-Harrison, J.S. (2000) The distribution, organization and evolution of two abundant and widespread repetitive DNA sequences in the genus *Hordeum*. Theor. Appl. Genet. 100, 169–176.
- Thompson, P.E. (1964) Evidence on the basis of the centromere effect in the large autosomes of *Drosophila melanogaster*. Genetics 49, 761–769.

- Van't Hof, J. and Sparrow, A.H. (1963) A relationship between DNA content, nuclear volume, and minimum mitotic cycle time. *Proc. Natl. Acad. Sci. USA* 49, 897–902.
- Van Schaik, N.W. and Brink, R.A. (1959) Transpositions of modulator, a component of the variegated pericarp allele in maize. *Genetics* 44, 725–738.
- Vershinin, A.V., Schwarzacher, T. and Heslop-Harrison, J.S. (1995) The large-scale genomic organization of repetitive DNA families at the telomeres of rye chromosomes. *Plant Cell* 7, 1823–1833.
- Vershinin, A.V. and Heslop-Harrison, J.S. (1998) Comparative analysis of the nucleosomal structure of rye, wheat and their relatives. *Plant Mol. Biol.* 36, 149–161.
- Vicient, C.M., Suoniemi, A., Anamthawat-Jónsson, K., Tanskanen, J., Beharav, A., Nevo, E. and Schulman, A.H. (1999) Retrotransposon BARE-1 and its role in genome evolution in the genus *Hordeum*. *Plant Cell* 11, 1769–1784.
- Vogel, K.P., Arumuganathan, R. and Jensen, K.B. (1999) Nuclear DNA content of perennial grasses of the Triticeae. *Crop Sci.* 39, 661–667.
- Werner, J.E., Kota, R.S., Gill, B.S. and Endo, T.R. (1992) Distribution of telomeric repeats and their role in the healing of broken chromosome ends in wheat. *Genome* 35, 844–848.
- Wicker, T., Stein, N., Albar, L., Feuillet, C., Schlagenhauf, E. and Keller, B. (2001) Analysis of a contiguous 211 kb sequence in diploid wheat (*Triticum monococcum* L.) reveals multiple mechanisms of genome evolution. *Plant J.* 26, 307–316.
- Wicker, T., Guyot, R., Yahiaoui, N. and Keller, B. (2003a) CACTA transposons in Triticeae. A diverse family of high-copy repetitive elements. *Plant Phys.* 132, 52–63.
- Wicker, T., Yahiaoui, N., Guyot, R., Schlagenhauf, E., Liu, Z.D., Dubcovsky, J. and Keller, B. (2003b) Rapid genome divergence at orthologous low molecular weight glutenin loci of the A and A^m genomes of wheat. *Plant Cell* 15, 1186–1197.
- Wicker, T., Zimmermann, W., Perovic, D., Paterson, A.H., Ganai, M., Graner, A. and Stein, N. (2005) A detailed look at 7 million years of genome evolution in a 439 kb contiguous sequence at the barley *Hv-eIF4E* locus: recombination, rearrangements and repeats. *Plant J.* 41, 184–194.
- Wicker, T., Sabot, F., Hua-Van, A., Bennetzen, J.L., Capy, P., Chalhoub, B., Flavell, A., Leroy, P., Morgante, M., Panaud, O., Paux, E., SanMiguel, P. and Schulman, A.H. (2007) A unified classification system for eukaryotic transposable elements. *Nat. Rev. Genet.* 8, 973–982.
- Xin, Z.-Y. and Appels, R. (1988) Occurrence of rye (*Secale cereale*) 350-family DNA sequences in *Agropyron* and other *Triticeae*. *Plant Syst. Evol.* 160, 65–76.
- Zhang, H.B. and Dvořák, J. (1990) Isolation of repeated DNA sequences from *Lophopyrum elongatum* for detection of *Lophopyrum* chromatin in wheat genomes. *Genome* 33, 283–293.

Chapter 24

Wheat and Barley Genome Sequencing

Kellye Eversole, Andreas Graner, and Nils Stein

Abstract A high quality reference genome sequence is a prerequisite resource for accessing any gene, driving genomics-based approaches to systems biology, and for efficient exploitation of natural and induced genetic diversity of an organism. Wheat and barley possess genomes of a size that was long presumed to be not amenable for whole genome sequencing. So far, only limited genomic sequencing of selected loci has been performed, providing preliminary information about the organization of the Triticeae genomes. Driven by breakthrough technology improvements, whole genome sequencing of Triticeae genomes is poised to become a realistic undertaking. This chapter provides an overview of the history of plant genome sequencing, summarizes the status of Triticeae genome sequencing efforts, describes next generation sequencing technologies, and offers an outlook on the future of wheat and barley genome sequencing based on these technologies.

24.1 Introduction

A genome sequence is an abrupt, rate-changing, transformative technology for genetics and all scientific disciplines relevant to understanding the biology of an organism. For every sequenced species, our knowledge and even the questions that can be asked about a genome changed as the sequence enabled a global perspective of the product of genotype and environment interactions, looking far beyond an individual gene or groups of genes for direct and efficient access to understanding biology. Ten years after obtaining the first eukaryotic genome sequence, it is routine to enumerate, in a single experiment, all the genes in an organism that respond to a specific stimulus or stress. For example, within 10 years of completing the yeast sequence, scientists were able to account for and model, in a fully quantitative way, not simply how each of the genes

K. Eversole (✉)
Eversole Associates, Bethesda, MD, USA
e-mail: eversole@eversoleassociates.com

participates in the expression of phenotypes but also how their interactions are controlled over the entire life cycle and range of environmental conditions in which they are able to exist (Dolinski and Botstein 2005).

The sequencing technique and assembly method selected for sequencing a particular genome depends on a number of factors, including genome size, complexity of the genome, cost, and, most importantly, the ultimate goals of the post-sequence research. So far, genome size and complexity have been the primary considerations for the selection of plant genomes to be sequenced without regard to the socio-economic relevance of the species. In this atmosphere, sequencing wheat and barley was out of scope for many years. However, the ongoing revolution in sequencing technology is resulting in significant reductions in cost and technological advancements in high throughput, thereby allowing us now to consider sequencing large, complex species such as those of the Triticeae tribe.

The importance of the Triticeae has been established well throughout this book; however, it cannot be overemphasized when one considers the methods and strategies for sequencing wheat and barley that the intended applications of the sequence should be the primary consideration when designing a process for sequencing the genomes of these two major cereal crops. Wheat, alone, is cultivated on more land than any other crop and, in some places, is replacing rice as the most important cereal for human consumption. Consumption of wheat and barley is increasing at a faster rate than the annual increases of yield growth resulting in depleted stocks and high prices. What does all this have to do with sequencing? Everything, as the sequences of these two genomes will be the foundation that supports breeder's efforts to accelerate improvement and ensure that the supply of these vital crops will be sufficient to meet the rising demand. If the foundation is not solid, the gains in time from using a rapid, but inaccurate sequencing method will pale in comparison to the loss in time for crop improvement.

A genome sequence is a powerful tool for understanding evolutionary processes and the underpinning biology of plants but its true power lies in the fact that it is the only platform that facilitates discovery of the entire regulatory network of developmental processes that result in complex traits, such as yield. We are not only interested in the differences between wheat and barley or wheat and rice or different varieties of any crop; rather, we are interested in how this information can be used for crop improvement, such as breaking through the existing yield barriers while maintaining essential nutritional qualities. As we learned with the Human, *Arabidopsis thaliana*, and rice genome sequences, the draft assemblies may have provided a majority of the genes but they were not adequate for understanding the complexity of gene regulation and expression that result in specific phenotypes. High quality reference genome sequences and comparative genomic studies were needed for complete gene discovery, accurate assembly, and analyses of complex traits in these species (Hanada et al. 2007; International Rice Genome Sequencing Project 2005; Olson 2008; Palmer and McCombie 2002; Sasaki et al. 2002; Wheeler et al. 2008). Since the

completion of high quality reference sequences for a number of species and the characterization of 1% of the human genome in the ENCODE project, it is becoming clear that regulation is dispersed across the majority of the genome and that noncoding intergenic regions, and “pseudogenes” are as important as the protein coding genes in regulating genome expression (Gerstein et al. 2007; Kidd et al. 2008; The ENCODE Project Consortium 2007; Thurman et al. 2007; Wilhelm et al. 2008; Yamada et al. 2003). Further, recent work has indicated that the vast majority of the mammalian genome is transcribed in a developmentally regulated manner (Pheasant and Mattick 2007). Thus, for wheat and barley, high quality reference sequences are essential for gaining access to the regulatory networks controlling complex agronomic traits.

The prospect of generating genome sequences for organisms with genomes as complex as those of barley (5 Gb) and wheat (17 Gb) is only possible as a result of the progress that has been made over the past 40 years. The first successful, albeit partial, sequence of an organism was published 40 years ago (Wu and Kaiser 1968) and it would take another two years to complete a 12-basepair sequence of the cohesive ends of the phage lambda (Wu and Taylor 1971). Modern chain-termination, dideoxynucleotide sequencing methods were introduced in 1977 by Sanger and colleagues. They were used initially to sequence the 5,386 bases of a small phage genome (Sanger et al. 1977) and have provided the foundation for automated fluorescent DNA sequencing that has been used most extensively for DNA sequencing since that time. The dideoxynucleotide sequencing method is based on the synthesis of labeled DNA strands that are complementary to random (shotgun) fragments of single-stranded DNA and are terminated at specific residues (Sanger 1988; Sanger 2001). Messing and his colleagues improved significantly the method by modifying the double-stranded bacteriophage M13 to contain a site into which DNA fragments to be sequenced and cloned can be inserted (Gronenborn and Messing 1978; Messing et al. 1977; Messing and Vieira 1982). Subsequent improvements introduced different labeling methods, the use of double-stranded DNA templates, automated fragment separation and detection, and strategies designed to sequence regions of DNA with complex secondary structure. From the time of the first complete genome sequence, almost 20 years would pass before the first nucleotide sequence of a eukaryotic genome would be completed – that of budding yeast (*Saccharomyces cerevisiae*) consisting of 12,068 kb (Goffeau et al. 1996).

The nematode *Caenorhabditis elegans* became the first multicellular organism to be sequenced and it paved the way for sequencing what were considered to be medium-size plant genomes – those between 30 and 300 Mb, as well as the human genome (~3 Gb). *C. elegans* was sequenced in parallel with budding yeast and used a similar strategy: relying on a clone-based physical map as a substrate, a tile-path of clones across the genome were each sequenced with a shotgun sequencing process and finished with directed sequencing to close gaps. Expressed Sequence Tag (EST) sequencing was used to identify and annotate the coding sequence (The *C. elegans* Sequencing Consortium 1998).

Following on from these early days of sequencing, significant advances in sequencing methods (e.g., improved chemistry, fluorescent dyes, etc.), throughput (e.g., automated capillary sequencers), and assembly software (e.g., alignment algorithms) as well as the development of complimentary sequencing strategies (e.g., whole genome shotgun combined with clone sequencing based on bacterial artificial chromosomes – i.e., BACs) enabled the sequencing of larger and more complex genomes and reduced sequencing costs substantially. Today, capillary-based Sanger sequencing can achieve read lengths of 1,000 bp with per-base accuracies as high as 99.9% at a cost of about \$0.50 per kilobase (Shendure and Ji 2008). However, despite the significant reductions of costs and increased rates of sequence data collection, *de novo* sequencing (i.e., sequencing a genome that had not been sequenced before) and assembly of higher plants and animal genomes is far from being a routine, simple, inexpensive process. By mid-2008, of the 97 eukaryotic genomes that have been sequenced with published results (<http://www.genomesonline.org>), only 14 are completely finished. Barley and wheat will follow a series of plant genomes to be sequenced and can benefit therefore from the experience of other projects and the modifications in strategies and advancements in sequencing technologies to obtain a high quality, reference genome sequence for a significantly reduced cost. To offer a context for the development of strategies for sequencing of the wheat and barley genomes, this chapter provides a brief history of sequencing in higher plants, the current status of sequencing efforts for the Triticeae genomes, and reviews the perspectives offered by the next generation DNA sequencing (NGS) technologies for sequencing these genomes in the near future.

24.2 History of Sequencing in Higher Plants

From the beginning of DNA sequencing efforts, the relatively large size and complexity of plant genomes have been major deterrents for consideration as a sequencing target. With the completion of the sequence of the yeast *Saccharomyces cerevisiae* in 1996, it was noted that sequencing becomes extremely complex and expensive for genomes greater than 6 Mb (Goffeau et al. 1996). Thus, while there was interest in sequencing economically important species, such as the maize genome, model species with small, diploid genomes were considered to be more tractable and important for contributing to knowledge of plant gene function in the first instance. To date of the 17 plant genomes with draft or completed sequences, all are diploid with a median size of 466 Mb (or an average of 687 Mb) (Table 24.1). Of the draft or completed genomes, only maize and soybean have a size greater than 1 Gb. Only eight higher plant genome projects have been completed and published and these represent just four species: rice (*Oryza sativa*), grapevine (*Vitis vinifera*), poplar (*Populus trichocarpa*), and thale cress (*Arabidopsis thaliana*).

Table 24.1 Overview of completed, draft, and ongoing higher plant genome sequencing projects

Species and genotype	Genome size (Mb)	Ploidy	Sequencing strategy	Coverage	References
Completed Sequencing Projects					
<i>Arabidopsis thaliana</i> (thale cress, cv. Columbia)	125	Diploid	BAC-by-BAC		www.arabidopsis.org
<i>Oriza sativa</i> ssp. <i>japonica</i> (rice, cv. Nipponbare)	389	Diploid	BAC-by-BAC	10X	http://rgp.dna.affrc.go.jp/IRGSP/
<i>Oriza sativa</i> ssp. <i>japonica</i> (rice, cv. Nipponbare by Syngenta)	433	Diploid	WGS	6X	Goff et al. (2002)
<i>Oriza sativa</i> ssp. <i>indica</i> (rice, cv. 93-11)	466	Diploid	WGS	6.28	http://rice.genomics.org.cn/rice/index2.jsp
<i>Oriza sativa</i> ssp. <i>japonica</i> (rice, cv. Nipponbare by Monsanto)	399	Diploid	BAC-by-BAC	5X	Barry (2001)
<i>Populus trichocarpa</i> (black cottonwood, poplar, cv. Nisqually-1)	485	Diploid	WGS	7.5X	http://genome.jgi-psf.org/Poptr1_1/Poptr1_1.home.html
<i>Vitis vinifera</i> (grapevine, Pinot Noir cv. PN40024)	487	Diploid	WGS	8.4X	www.vitaceae.org
<i>Vitis vinifera</i> (grapevine, Pinot Noir cv. ENTAV 115)	505	Diploid	WGS & SBS	6.5X + 4.2X	http://genomics.research.iasma.it/iasma/
Draft Genome Sequences Available					
<i>Brachypodium distachyon</i> (purple false brome, line Bd21)	300	Diploid	WGS	8X	http://www.brachypodium.org/
<i>Carica papaya</i> (transgenic papaya, “SunUp”)	372	Diploid	WGS	3X	Ming et al. (2008)
<i>Glycine max</i> (soybean, cv. Williams 82)	1,100	Diploid	WGS	~7X	http://www.phytozome.net/soybean

Table 24.1 (continued)

Species and genotype	Genome size (Mb)	Ploidy	Sequencing strategy	Coverage	References
<i>Lotus japonicus</i> (trefoil, cv. Miyakojima MG-20)	472	Diploid	BAC-by-BAC	8.4X	http://www.kazusa.or.jp/lotus/
<i>Minulus guttatus</i> (common monkeyflower)	430	Diploid	WGS	7X	http://www.jgi.doe.gov/genome-projects/
<i>Ricinus communis</i> (castor bean, cv. Hale)	400	Diploid	WGS	4X	http://www.jvri.org/cms/research/projects
<i>Sorghum bicolor</i> (sorghum, cv. BTx623)	770	Diploid	WGS	8X	http://www.phytozome.net
<i>Zea mays</i> ssp. <i>mays</i> (maize, cv. B73)	2,600	Diploid	BAC-by-BAC (gene space)	6X	http://maizesequence.org
<i>Zea mays</i> (popcorn, cv. Palomero Toluqueno)	2,100	Diploid	Hybrid WGS (10% HiCot/90% MF) + pyrosequencing (50% MF, 50% WGS)	3.2X	http://www.maizegdb.org/cgi-bin/displayrefrecord.cgi?id=1188121
Sequencing Projects in Progress or Not Publicly Available					
<i>Aquilegia formosa</i> (Western columbine)	400	Diploid	WGS		http://www.jgi.doe.gov/genome-projects/
<i>Arabidopsis lyrata</i> (rock cress)	230	Diploid	WGS		http://www.jgi.doe.gov/genome-projects/
<i>Brassica oleracea</i>	600	Diploid	WGS		http://www.genomesonline.org
<i>Brassica rapa</i> ssp. <i>pekinensis</i> (Chinese cabbage, cv. Chiifu 401–42)	529	Diploid	BAC-by-BAC		http://www.brassica.info
<i>Capsella rubella</i> (pink shepherds purse)	250	Diploid	WGS		http://www.jgi.doe.gov/genome-projects/
<i>Citrus sinensis</i> (sweet orange, cv. Ridge Pineapple)	382	Diploid	WGS		http://www.citrusgenome.ucr.edu/
<i>Elaeis guineensis</i> (oil palm, cv. tenera x dura)	1,800	Diploid	unpublished		www.syntheticgenomics.com/press/2008-05-21.htm

Table 24.1 (continued)

Species and genotype	Genome size (Mb)	Ploidy	Sequencing strategy	Coverage	References
<i>Eucalyptus grandis</i> (BRASUZ1)	600	Diploid	WGS		http://www.jgi.doe.gov/genome-projects/
<i>Gossypium raimondii</i> (cotton)	880	Diploid	WGS		http://www.jgi.doe.gov/genome-projects/
<i>Jatropha curcas</i> (Synthetic Genetics)	500	Diploid	unpublished		http://www.syntheticgenomics.com/press/2008-05-21.htm
<i>Manihot esculenta</i> (cassava, cv. AM560-2)	770	Amphiploid	WGS		http://www.jgi.doe.gov/genome-projects/
<i>Medicago truncatula</i> (barrel medic, cv. Jemalong A17)	500	Diploid	BAC-by-BAC		http://www.medicago.org/
<i>Nicotiana tabacum</i> (tobacco, cv. Hicks Broadleaf)	4,500	Tetraploid	Methyl filtration		http://www.tobaccongenome.org/
<i>Setaria italica</i> (foxtail millet)	515	Diploid	WGS		http://www.jgi.doe.gov/genome-projects/
<i>Solanum lycopersicum</i> (common tomato, cv. <i>esculentum</i> × <i>pennellii</i>)	950	Diploid	BAC-by-BAC (gene space)		http://sgn.cornell.edu/about/tomato_project_overview.pl
<i>Solanum tuberosum</i> (potato, RH89-039-16)	840	Diploid	BAC-by-BAC		www.potatogenome.net
<i>Vigna unguiculata</i> (cowpea)	620	Diploid	Methyl filtration		Timko et al. (2008)
<i>Zea mays</i> (maize, cv. Mo17)	2,100	Diploid	WGS		http://www.maizegdb.org/sequencing_project.php

24.2.1 The First Plant Genome Model – *Arabidopsis thaliana*

The first and, to date, the most accurate sequence of a plant is that of thale cress, *Arabidopsis thaliana*. During the mid-to-late 1980s, interest in using *Arabidopsis* as a model system for plant genomics grew because of its small genome size (~150 Mb), short generation time, and large number of offspring. After a 1989 meeting to explore the utility of *A. thaliana* as a model, national and international steering committees were formed to develop a series of genomic resources that would culminate ultimately in the sequencing of the *Arabidopsis* genome (Somerville and Koornneef 2002). The *A. thaliana* sequence was completed to a level enabling gene annotation and publication in 2000 and finished to high quality subsequently. The strategy chosen to sequence the genome was a clone-based shotgun sequencing approach similar to that used for the model *C. elegans* for sequencing and finishing large, contiguous DNA fragments. The substrates for sequencing were BAC clones constructed from DNA from the accession Columbia. Sets of minimally overlapping BACs were selected from physical maps assembled on the basis of restriction fragment “fingerprints” and alignment with genetic markers that placed clone contigs on chromosomes (<http://nucleus/cshl.org/arabmaps/>). End sequences (http://www.tigr.org/tdb/at/abe/bac_end_search.html) of 47,788 BAC clones were used to extend and integrate contigs from BACs anchored by marker content. The total length of the sequenced regions, which extend from either the telomeres or ribosomal DNA repeats to the 180-base-pair (bp) centromeric repeats, is 115,409,949 bp. Estimates of the unsequenced, centromeric and rDNA repeat regions measure roughly 10 megabases (Mb), yielding a genome size of about 125 Mb, in the range of the 50–150 Mb haploid content estimated by different methods (Table 24.1).

The annotation of the *A. thaliana* genome sequence and analysis of gene function has continued since the release of the first assembly. With each new release of The Arabidopsis Information Resource (TAIR) genome annotation, the number of protein coding genes has increased. TAIR8, released in 2008, contains 27,235 protein coding genes, 4,759 pseudogenes or transposable elements, and 1,288 non coding RNAs (33,282 genes in all, 38,963 gene models) adding a total of 1,291 new genes and 2009 new gene models to the previous release (<http://arabidopsis.org>, Swarbreck et al. 2008). Efforts are underway now to sequence 1,001 *Arabidopsis* accessions with new sequencing technologies to determine the sequence variation that gives rise to the wide range of adaptations to habitat exhibited by this plant species (<http://1001genomes.org/>).

24.2.2 The First Economically Important Plant Genome – Rice

By the late 1990s, rice, possessing the smallest genome (approximately 400 Mb) of the major cereals, was poised to enter the whole genome sequencing queue.

With a high density physical map in hand, international support grew to use rice as a model for sequencing cereal genomes. After several meetings in 1997 and the submission by a Japanese group of the first proposal to sequence rice, a workshop was held in Singapore in September 1997 to explore the feasibility of an international, collaborative effort and to establish an international rice genome working group. Although the genome was almost four times the size of *A. thaliana*, it was considered feasible and affordable to sequence a 400 Mb genome. Thus, in January, 1998, the Rice Genome Project in Japan was awarded funding to sequence the genome and subsequent countries secured funding for participation in what became the International Rice Genome Sequencing Project (Burr 1999; Messing 1999; Sasaki and Burr 1999).

The international group used DNA from a single rice variety out of concerns that allelic polymorphisms could hamper accurate compilation of the sequence. The group selected *Orzya sativa* L. ssp *japonica* cultivar “Nipponbare” for sequencing as it had been used for extensive EST sequencing. A whole genome shotgun sequencing approach (i.e., sequencing random fragments of genomic DNA rather than pre-mapped bacterial clones) was considered but was not selected as there were concerns about the accuracy and completeness achievable with such an approach. Thus, a clone-by-clone strategy using BACs and PACs with an aim of achieving 99.99% accuracy was followed (Messing 1999; Sasaki and Burr 2000). An interim goal for the international group was to achieve draft quality (i.e., sequenced BACs or P1 artificial chromosomes – PACs – with few sequencing gaps whose pieces are ordered and oriented directionally) by the end of 2002 (Burr 2002).

The first rice genome sequence released was not that of the international effort; rather, Monsanto released a draft sequence of the same cultivar (Nipponbare) in 2001. BACs were sequenced using a clone-by-clone approach to the level of 5X coverage and assembled into a draft data set comprising 259 Mb of sequence data (Barry 2001; Vij et al. 2006). At the same time, Syngenta (Torrey Mesa Research Institute) joined with Myriad Genetics to sequence the same cultivar using a whole genome shotgun approach and achieved an assembled sequence that covered 93% of the genome with more than 6-fold coverage (Goff et al. 2002). Also utilizing a whole genome shotgun approach, the Beijing Genome Institute (BGI) sequenced 93-11, a cultivar of the *indica* subspecies to a 4.2X coverage, initially, and to 6.28X subsequently. Including all unmapped fragments and unassembled reads, the total genome size is 466.3 Mb for Beijing *indica* (Yu et al. 2002, 2005) (Table 24.1).

With the completion of the four draft sequences by 2002, a debate ensued as to whether draft sequences provided an adequate resource on which to base crop genetics research. The limitation of the draft sequences were the assembly and base quality errors inherent in a low coverage sequence and the incomplete coverage of the genome (Buell 2002). When the draft genomes were compared with the complete sequences of chromosomes 1 and 4 generated by the International Consortium, it became clear that a high quality finished BAC-by-BAC sequence would be the only means of obtaining an appropriate reference

genome for studies of gene function and regulation and for rice genetics research since 43% of the genes predicted from whole-genome shotgun sequence methods were incomplete and many were not associated with the essential regulatory sequences that fall outside of transcribed regions (Sasaki et al. 2002).

Remaining committed to obtaining a high quality finished genome, the scientists involved in the International Rice Genome Sequencing Project reached this goal in 2005 using YAC-and BAC-based physical maps and additional information from a high density genetic map, ESTs, BAC-end sequences, and the two draft genome sequences (International Rice Genome Sequencing Project 2005). The sequence achieved “finished” quality with less than one error in 10,000 nucleotides, having resolved all ambiguities, and attempted all state-of-the-art efforts to close gaps. A total of 3,401 BAC/PAC clones were sequenced to 10X coverage, assembled and ordered, resulting in 370 Mb of finished sequence that represented 95% of the genome and virtually all euchromatic regions (Table 24.1). The sequence revealed a genome size of 389 Mb, slightly lower than the estimated 400 Mb genome size (International Rice Genome Sequencing Project 2005). Through the Rice Annotation Project, the sequence first was annotated automatically and then manually curated. The genome has approximately 32,000 genes with repetitive elements comprising 30% of the genome (Itoh et al. 2007).

24.2.3 The First Tree Genome – Poplar Genome Sequence

In contrast to the rice and *A. thaliana* sequencing projects, an international consortium was not established to coordinate the sequencing of the poplar genome; instead, sequencing was conducted by the Joint Genome Institute (JGI) of the US Department of Energy and the International *Populus* Genome Consortium was established to assist in post-sequence activities, including genome annotation (<http://www.ornl.gov/sci/ipgc/home.htm>). The genome of the highly heterozygous black cottonwood tree, *Populus trichocarpa*, was selected by the JGI as a model for sequencing forest species because of its small size (~485 Mb or 50 times smaller than pine (<http://www.jgi.doe.gov/poplar/>)). A single genotype, Nisqually-1, was sequenced to 7.5X coverage using a whole-genome shotgun strategy (Table 24.1). Although a BAC based physical map was used to guide the sequence assembly, it was possible to assemble only the euchromatic region of the genome, leaving a substantial portion (~75 Mb) of the whole genome shotgun reads unassembled. A significant fraction of the unassembled reads was repetitive DNA or possible repetitive DNA that had not yet been characterized in *Populus* or other organisms. Annotation of the 410 Mb of assembled scaffolds indicated an initial set of 45,555 protein-coding genes. BAC ends were sequenced to assist assembly of the scaffold and to anchor the physical map (Kelleher et al. 2007; Tuskan et al. 2006).

Community-wide, manual curation of predicted genes is ongoing (http://genome.jgi-psf.org/Poptr1_1/Poptr1_1.info.html).

24.2.4 Two Grapevine Sequences

Genomic efforts on grapevine began in the mid-1990s. The small size of the *Vitis vinifera* genome (~475 Mb) and the successful completion of the *A. thaliana* and rice genomes led to the establishment of the International Grape Genome Project in 2001 as a vehicle to advocate for sequencing a grapevine cultivar, commonly called “vinifera” or “Grape” (<http://www.genoscope.cns.fr/spip/Vitis-vinifera-whole-genome.html>). In 2005, INRA and Genoscope (both of France) joined with several Italian groups to form “Vigna/Vigne”, the French-Italian Public Consortium for Grapevine Genome Characterization, to sequence the grapevine genome (Travis 2008). As most vinifera grapevines are extremely heterozygous, the public consortium selected a Pinot Noir variety, PN40024, with a simplified genome bred to almost full homozygosity (~93%) rather than a variety used for wine production. Genome coverage of 8.4-fold was achieved by whole genome shotgun sequencing and 487.1 Mb were assembled (Table 24.1). The sequence revealed 30,434 protein-coding genes and a genome containing approximately 41% repetitive or transposable elements (The French–Italian Public Consortium for Grapevine Genome Characterization 2007).

At the same time that the public consortium was working on the highly inbred variety, the Istituto Agrario di San Michele All’Adige (IASMA) launched a private effort to sequence the heterozygous Pinot Noir accession ENTAV 115 that is cultivated for wine production. To decode the heterozygous genome, a hybrid approach was utilized whereby 6.5-fold coverage in whole genome shotgun sequence generated on automated capillary sequences (Sanger sequencing) was integrated with 4.2X coverage in sequence reads generated using the Roche-454 highly parallel sequence by synthesis (SBS) method (see below for a description of the technology) (Table 24.1). Approximately 477 Mb were assembled. The genome size was estimated to be 504.6 Mb with approximately 27.4% repetitive regions and 29,585 predicted genes (Velasco et al. 2007).

24.2.5 The First Moderately-Sized Plant Genome Sequence – Maize

Even though the maize genome sequence has not been published or annotated as of mid-2008, it is worthwhile to review here the sequencing strategy used for this relatively complex genome as it is the first moderately-sized plant genome to be sequenced. Further, rice, wheat, maize, and barley are the most important

cereal crops and the quality of the genome sequences for each of these will affect efforts to accelerate crop improvements. The maize genome is smaller in size (~2.3 Mb) than barley (~5 Gb) or wheat (~17 Gb), however, it is much more complex than any of the other plant genomes sequenced to date. Comparable to the wheat and barley genomes, the maize genome has a high percentage of repetitive elements, although recent results suggest that the repetitive content of maize may be as low as 58% (Messing et al. 2004).

While significant research was conducted to develop a strategy for sequencing the maize genome as well as to build numerous genomic resources between the mid-1990s and 2005, the focus here will be on the method selected for sequencing. Beginning in 2005, a multi-institutional collaboration led by the Genome Sequencing Center at Washington University, St. Louis, was funded by the US National Science Foundation, the US Department of Agriculture, and the US Department of Energy to sequence the *Zea mays* ssp. *mays* cv. B73.

An approach based on sequencing BACs selected from the sequence-ready, fingerprinted contig-based maize physical map (Wei et al. 2008) was utilized to achieve high-quality (fewer than one error per 100,000 bases) sequence coverage of the non-repetitive, identifiable gene containing regions (Wilson et al. 2007) (<http://genome.wustl.edu>; www.maizesequence.org/overview.html). The first draft of the genome was released in February 2008 based on a “freeze date” of December 2007. The latest release 2a.50 (June 2008 freeze date), includes 16,205 BACs with total DNA of 2.7 Gb, and 57,480 protein-coding genes (<http://www.maizesequence.org>) (Table 24.1). Another release is expected in late 2008. Analysis of the sequence is underway. It is possible that the minimal tile path of BACs selected for sequencing could miss 5–25% of the maize genome (Rabinowicz and Bennetzen 2006) but it is too early to assess fully the extent of the coverage of the genome.

To complement the BAC-by-BAC approach, shotgun sequencing with next generation sequencing (NGS) technologies is in progress for the gene-space of *Zea mays* ssp. *mays* cv. Mo17. Using Roche-454-FLX sequencing, up to 12-fold coverage of the genome will be achieved with paired-end sequence reads of 100–200 bases in length which is expected to identify short-range structural variation and ultra-short read Illumina or ABI-SoLID sequence data will be used for polishing. Assembly and annotation is underway (http://www.maizegdb.org/sequencing_project.php).

24.2.6 Other Plant Genome Projects

With advancements in sequencing technology and declining costs, more and more plant genome sequencing projects are getting underway. The status of most of the *de novo* sequencing projects for the genomes of higher plants is provided in Table 24.1 with 7 genomes for which drafts are available and another 18 that are in progress or are not yet publicly available.

24.3 Current Status of Triticeae Genome Sequencing

Genome size and sequencing costs have so far impeded a systematic sequencing of the wheat and barley – and the Triticeae genomes in general. Nevertheless, insights into genome organization and complexity at the sequence level have been gained by a number of complementary approaches (Stein 2007). As has been done for many other species when full genome sequencing is out of reach, random genomic information was accumulated for wheat and barley by sequencing either expressed genes from cDNA libraries (expressed sequence tags, EST), genomic DNA from BAC insert ends (BAC end sequencing, BES) or genomic libraries enriched for genic sequences (Methyl-filtration, MF; High- C_{ot} filtration, CF) (Table 24.2). Larger, contiguous genomic sequence information was obtained mainly by sequencing individual BAC clones or small BAC contigs as part of map-based isolation of agronomically important genes. A brief overview of the sequence resources available to date for the Triticeae genomes is provided below.

Table 24.2 Status of genome sequencing in Triticeae species

	ESTs ¹	GSS ²	genomic sequences (>70,000 bp) ³	
<i>Triticum aestivum</i>	1,051,279	45,506	30.0 Mb	248 BACs
<i>Triticum monococcum</i> (<i>urartu</i>)	10,139	37(3)	2.5 Mb	15 BACs
<i>Triticum turgidum</i> ssp. <i>durum</i> (<i>dicoccoides</i>)	19,319	5 (4)	4.4 Mb	24 BACs
<i>Hordeum vulgare</i> ssp. <i>vulgare</i>	478,682	2,033	3.9 Mb	23 BACs ⁴
<i>Hordeum vulgare</i> ssp. <i>spontaneum</i>	24,161	–	–	–
<i>Secale cereale</i>	9,293	81	–	–
<i>Aegilops tauschii</i>	116	5,055	1.0 Mb	7 BACs

¹ Expressed Sequence Tags, dbEST release 032108 (http://www.ncbi.nlm.nih.gov/dbEST/dbEST_summary.html)

² Genomic Survey Sequences, dbGSS release 032108 (http://www.ncbi.nlm.nih.gov/dbGSS/dbGSS_summary.html)

³ Surveyed at NCBI <http://www.ncbi.nlm.nih.gov/sites/entrez> by March 28, 2008

⁴ Includes contigs consisting of more than a single BAC, Table modified and updated from Stein (2007)

24.3.1 EST Sequencing

Accumulation of expressed sequence tags (EST) – short sequence reads obtained from expressed genes (cloned cDNA) – is a straightforward and less costly way to access large portions of the gene content of a species (Rudd 2003). The comprehensiveness or completeness of the gathered gene content information can be evaluated by direct comparison to completely sequenced reference or model genomes, such as those of *A. thaliana* or rice – an approach known as “reconstructomics” (Rudd 2003) since the putative gene content of a

non-sequenced genome is being reconstructed on the basis of a model genome. Almost 500,000 barley ESTs and over a million wheat ESTs have been produced mostly by the International Triticeae Mapping Initiative community (ITEC, <http://wheat.pw.usda.gov/genome>) and are stored in public databases for barley and wheat (Table 24.2, Chapter 9), respectively. Independent assemblies of different proportions of these public datasets have been performed and revealed alternative, so called “unigene” sets (unique gene consensus sequences = independent EST contigs that tentatively represent independent genes) (Lazo et al. 2004; Ogihara et al. 2003; Zhang et al. 2004 and Table 24.3). The wheat and barley EST resources were instrumental for designing standardized tools for transcript profiling based on the Affymetrix Genechip technology (Bhat et al. 2007; Close et al. 2004) and were utilized intensely as resources for high-throughput and high-density meiotic mapping in barley or physical assignment to linkage groups in bread wheat, respectively (Qi et al. 2004; Rostoks et al. 2005; Stein et al. 2007).

Table 24.3 Overview of tentative “unigene” assemblies for wheat and barley

Unigene assemblies	Wheat			Barley		
	# ESTs	Contigs	Singles	# ESTs	Contigs	Singles
TIGR ¹	580,155	44,954	77,187	370,546	23,176	27,094
HarvEST ²	101,107	16,413	20,616	444,652	28,001	22,937
NCBI unigene ³	947,177	32,900	8,326	456,769	21,376	1,702

¹ <http://compbio.dfci.harvard.edu/tgi/plant.html>, wheat = assembly 10.0, January 14, 2004; barley = assembly 9.0 September 15, 2004

² <http://www.harvest-web.org/>, wheat = HarvEST:Wheat version 1.14, November 26, 2007, barley = HarvEST:Barley version 1.65, March 29, 2008

³ <http://www.ncbi.nlm.nih.gov/sites/entrez?db=unigene>, wheat = unigene build #51, January 24, 2008, barley = unigene build #52 January 30, 2008

24.3.2 GSS

Genomic survey sequencing (GSS) is the attempt to harness a more or less random and representative survey of a genome in order to deduce aspects of genome organization, complexity, and repetitiveness. Shotgun sequencing of about 1-fold haploid genome coverage reveals representative information about repetitive DNA and gene content and anchoring these sequences to a reference genome can help to deduce the level of genome collinearity and sequence conservation. Until recently, to cover a diploid, Triticeae genome (~5 Gbp) by 1-fold Sanger shotgun sequencing, at a cost of \$0.50 (USD) and an average length of 700 bp per read, a \$4 million (USD) investment would be required. This high cost has been the principal reason why such an approach, based on classical Sanger technology, has not been applied yet to wheat and barley. This situation has changed as the cost of sequencing using Sanger

methods has been cut in half and as a result of significant innovations in sequence technology (see below).

An alternative approach to accumulate practically un-biased genomic information is to generate sequence reads from the ends of BAC clones (BES) in concert with the development of a BAC-based physical map of a genome. Here, end sequences of large insert clones (BES) that are analysed by high-throughput BAC clone fingerprinting during physical map construction (see Chapter 11) are sequenced providing information for marker development and direct anchoring of physical BAC contigs via sequence identity to already mapped markers or sequenced BAC clones (sequence-tagged connector strategy, STC, Venter et al. 1996). In the framework of developing a physical map for wheat chromosome 3B, a total of 19,399 BES were obtained by paired-end sequencing of 21,504 BACs (Paux et al. 2006). These data indicated that perhaps more than 95% of the wheat genome consists of non-genic, mainly repetitive sequences – important information for any attempt at *de novo*, full-genome sequencing of such a large and complex genome. As part of the project to establish and anchor a physical map of the barley genome, paired-end sequencing of 250,000 BACs is in progress (Robbie Waugh, Michele Morgante, and Nils Stein, unpublished data). This will produce roughly 350 Mb of single pass, random genomic sequence information representing about 5% of the barley genome.

A number of different technologies were employed to preferentially sequence the genic regions from large, complex, and highly repetitive genomes. Methyl-filtration (MF, Rabinowicz 2003) and hypomethylated partial restriction cloning combined with sequencing (HMPR, Emberton et al. 2005) take advantage of the fact that gene-bearing DNA is less methylated (i.e., hypo-methylated) in plants compared to non-coding/repetitive DNA. Another approach to enrich for gene sequences in genomic libraries is based on slower re-association kinetics of low-copy non-repetitive DNA upon complete denaturation of double stranded- genomic DNA (High- C_{ot} , cot filtration, CF, Peterson et al. 2002). None of the above mentioned approaches is perfect regarding the enrichment for genic vs. non-genic sequences; however, substantial enrichment (13–18 fold) for gene derived sequences was obtained by CF in wheat (Lamoureux et al. 2005) and MF in barley (Rabinowicz et al. 2005), respectively. Thus, gene-enriched sequencing can be considered a relevant complementary task to survey the gene space of the complex genomes of wheat and barley.

24.3.3 *Contiguous Genomic DNA Sequences*

Larger contiguous sequences of wheat and barley genomic DNA have been obtained so far mainly within projects aiming at map-based isolation of genes (Stein 2007). Although the limited dataset was accumulated from about a dozen selected loci only, it provided useful information about genome structure and organization (e.g. Anderson et al. 2003; Dubcovsky et al. 2001; Rostoks et al.

2002; Wicker et al. 2001; Wicker et al. 2005; Yan et al. 2003), but it is, by far, not representative of the entire wheat and barley genomes. This is indicated by the sequences containing a gene density that is estimated to be a factor of 2–5 higher than the theoretical and extrapolated gene density based on a haploid Triticeae core genome size of 5 Gb and an overall gene content of between 30 and 50,000 genes (Stein 2007). Recently, a project was initiated to sequence 220 randomly selected BAC clones of hexaploid wheat in order to obtain a less biased view of wheat genome organization. A preliminary analysis of four randomly selected clones revealed an average gene density of 1 gene in 75 kb. Since three of the clones were thought to originate from “gene-rich” regions of the wheat genome, this finding suggests that genes might be more evenly distributed throughout the genome as opposed to forming larger clusters or “gene islands” (Devos et al. 2005).

Sequencing of larger contiguous regions of the genome will require efficient tools for sequence annotation to inform the process of sequence and genome assembly. Analysis of the nested organization of wheat and barley repetitive elements led to the construction of a dedicated database (<http://wheat.pw.usda.gov/ITMI/Repeats/>) for the repeat element families of the wheat and barley genomes (Wicker et al. 2002). It also provided insights into genome organization even in stretches of DNA that extend over several 100 kb and do not contain any or only few genes (Komatsuda et al. 2007; Wicker et al. 2001), and also helped to assess the evolutionary time-span in which the repetitive landscape of the Triticeae genomes has been reshaped substantially (Ramakrishna et al. 2002; Wicker et al. 2003; Wicker et al. 2005).

24.4 Next Generation Sequencing (NGS) Technologies

Stimulated by a request for applications (RFA-HG-04-003) by the US American National Institute of Health (NIH), “REVOLUTIONARY GENOME SEQUENCING TECHNOLOGIES – THE \$1000 GENOME” (published February 12, 2004), several new technology platforms have been produced that achieve significant reductions in costs of sequence generation as well as faster and higher sequencing throughput (Mardis 2008 and Table 24.4) with the ultimate

Table 24.4 Comparison of estimated costs of NGS technology¹

Technology	bp/run	Costs ² /Mb	Time of run/Gb ³
Sanger	70 kb	\$2,000	~18,000 h
454/Roche GSFLX	100 Mb	\$60	~80 h
Illumina Genome Analyzer	1 Gb	\$2	~96 h
ABI SOLID	3 Gb	\$2	~192 h
HeliScope Helicos	> 1 Gb	\$1	~24 h

¹ Excluding purchase of equipment

² According to Shendure and Ji (2008) and <http://www.helicosbio.com>

³ Plain sequencing run time

goal of making re-sequencing of human genomes a central and routine component of medical diagnostics. These new platforms are referred to as “second generation” or “next generation” sequencing (NGS) technology since they no longer rely on classical, gold-standard “Sanger” sequencing methodology (Sanger et al. 1977).

Several articles summarize the different facets of these new technological developments (Bennett et al. 2005; Hutchison 2007; Metzker 2005; Ryan et al. 2007; Service 2006; Shendure and Ji 2008; Shendure et al. 2004). Here, we focus on systems that have entered the market already, such as Roche/454 GSFLX, Illumina Genome Analyzer, Applied Biosystems SOLID systems, and the HeliScope of Helicos with respect to their potential for sequencing the Triticeae genomes.

24.4.1 *Roche-454 GSFLX*

The first NGS system, launched commercially in 2004, is based on highly parallelized pyrosequencing technology (Ronaghi 2001). A pyrophosphate (PPi) molecule is released when a deoxy-nucleotide (dNTP) is incorporated into a DNA strand. Subsequently, PPi is converted enzymatically to ATP which provides energy for oxidation of the molecule Luciferin to Oxyluciferin – a process leading to the emission of light that can be detected by current CCD-camera technology. The mode of dNTP dispensing and signal occurrence/signal intensity can be translated into a nucleotide sequence. The Roche-454 technology applies pyrosequencing – or sequencing by synthesis (SBS) – in a highly parallelized manner (Margulies et al. 2005). Double-stranded DNA molecules are ligated to adaptors that allow subsequent binding to small beads – ideally at an optimal ratio of one DNA molecule per bead. Emulsion PCR is used then to clonally amplify the DNA on the beads so that, following removal of one of the strands and initiating DNA synthesis with DNA polymerase and deoxribonucleotides, a detectable signal is generated when a single nucleotide is incorporated into the growing DNA strand and hydrolysis of the pyrophosphate molecules released are coupled to a light producing luciferase reaction. The intensity of the light emission is proportional to the number of nucleotides incorporated and, therefore, runs of the same nucleotide are represented by a higher signal but are difficult to de-convolute if the number of bases is greater than 5 or 6. In highly repetitive genomes, this feature can lead to reduced sequence accuracy in specific regions. For the sequencing procedure, beads are spread into so-called “picotiterplates” – arrays formed by cut, optic fibre bundles where each fiber provides a single cavern large enough to accommodate only a single bead of the emulsion PCR. Light signal sequences are recorded for each individual bead leading to the parallelized recording of several hundred thousands of individual sequence reads. The first system launched in 2004 (GS20), was capable of recording 20 Mb of sequence comprising reads with

an average length of 100 nucleotides (nt) in a single 4 h run. This was upgraded in 2007 (GSFLX) to yield 100 Mb of ~250 nt reads in 8 h (Table 24.3). A further improvement has been announced for 2008 (Titanium) to allow sequencing of 0.5 Gb of >400 nt reads at costs similar to the current system (Table 24.4). This rapid technology advancement was achieved by improving enzyme and reaction chemistry, image capture technology, and image analysis algorithms while maintaining the basic principle of pyrosequencing.

24.4.2 Illumina Genome Analyzer

The Illumina Genome Analyzer, launched in 2006, produces about 1 Gb of sequence information from a single run at significantly shorter read lengths of up to 35 nt (Table 24.4). Randomly fragmented genomic DNA is fused to linker-adaptors allowing covalent attachment to a glass-slide like carrier called “flow-cell” (Bennett et al. 2005). The adaptors serve also as priming sites for subsequent solid-phase PCR that leads to clonal amplification and local cluster formation of the originally attached templates. Over 50 million of such well separated clusters, each containing ~1,000 copies of the same molecule, can be formed in the same flow cell. Similar to pyrosequencing, sequence determination occurs by synthesis and detection of the incorporation of fluorescently labeled nucleotides blocked or inactivated for 3' strand extension. Incorporation of labeled nucleotides results in fluorescently labeled clusters. After each round of strand extension, image capture records the fluorescence emission from each cluster which is translated into nt sequences. De-blocking of the 3' end and removal of the fluorescent signal of incorporated nucleotides prepares for the next step of strand extension.

24.4.3 Applied Biosystems SOLiD (Sequencing by Oligo Ligation and Detection)

The SOLiD platform uses a process of sequencing by ligation. Fluorescently labeled octamers are ligated to clonal DNA molecules that are amplified on small diameter beads using emulsion PCR and distributed onto glass slides (Shendure et al. 2004). A limited set of semi-degenerated 8-mer oligo-nucleotides that differ either at their 4th and 5th or 1st and 2nd nt-position – which determines the fluorescent label of the respective oligo-nucleotide – is dispensed automatically and the signal recorded. Cleavage of the ligated oligos at the 6th nt position leads to release of the fluorescent label and prepares for the next oligo-nucleotide ligation (for further detail see <http://solid.appliedbiosystems.com>, and Mardis 2008; Shendure and Ji 2008). The system currently produces in an eight day run about 3 Gb of sequence data (Table 24.4) consisting of short paired reads of 35 nt

length at a level of accuracy that is improved through the recording of two independent signals for each base sequenced.

24.4.4 HeliScope, Helicos

The HeliScope is currently the only NGS system available that directly sequences single DNA molecules and does not require prior amplification (Harris et al. 2008). A highly sensitive fluorescent imaging system is able to detect the incorporation of single nucleotides into growing DNA strands that are tethered to a slide surface. Fluorescence incorporation is measured for each of the millions of tethered strands on separate addition of each of the four deoxynucleotides of DNA and the fluorescent labels are, subsequently, cleaved before addition of the next nucleotide. The asynchronous strand extension is translated into nucleotide sequences that typically are in the range of 25–35 bp (more details are provided at <http://www.helicosbio.com>) The system has principally been designed to deliver over 1 Gb of sequence data in less than 24 h (Table 24.4).

24.4.5 Impact on Triticeae Genome Sequencing

Altogether, NGS technologies provide significantly higher sequencing throughput at lower costs than classical Sanger sequencing with dye-terminator chemistry and capillary sequencing devices (Table 24.4). They, therefore, satisfy a key prerequisite for sequencing several gigabase-sized crop plant genomes, such as those of wheat and barley. Further innovative sequencing technology concepts that build on direct sequencing of single molecules (“third generation” sequencing technology) are anticipated shortly (Hardin et al. 2004; Harris et al. 2008; Korlach et al. 2008, for a review see Ryan et al. 2007; Shendure and Ji 2008) and promise that the increase of sequencing throughput and cost reduction will continue at the current pace. With this perspective and especially for large and complex genomes, the true challenges in the future likely will not be in generating sufficient sequence information in predictable time frames, but rather to provide sufficient storage capacity and calculation power for handling the massive quantities of data that will accumulate over short periods of time in the near future (Hutchison 2007; Pop and Salzberg 2008).

All available NGS methods principally provide sufficient capacity to generate a ten-fold whole genome shotgun sequence dataset for a 5 Gb Triticeae genome like barley on a single apparatus within a timeframe of 1–12 months. But, as mentioned in the introduction, will this be sufficient to obtain the quality of sequence needed for the wheat and barley genomes? In other words, can we assemble the 25–50 nt sequences obtained by Illumina/Solexa, ABI/SOLiD HeliScope/Helicos, and the 250 nt (soon 400 nt) currently obtained

by GSFLX from Roche-454 into a complete sequence? The short but highly accurate reads of the initial three platforms were designed for and are highly instrumental for re-sequencing known genomes in order to identify variations in individuals (Warren et al. 2008). This approach is applicable in plants only for the *A. thaliana* and rice genomes (i.e., Cokus et al. 2008) where high-quality reference sequences are available. For new genomes, the short read lengths make *de novo* sequence assemblies possible only for small genomes void of repetitive DNA (Chaisson and Pevzner 2008; Hernandez et al. 2008). However, in combination with paired-end or mate-pair sequencing, short-read NGS technology has a strong potential to become a powerful component of *de novo* sequencing in complex plant genomes as well (Shendure and Ji 2008), given that the paired-end “libraries” are of sufficient size to bridge repeat elements in the Triticeae genomes. The potential of 454 pyrosequencing for *de novo* sequencing genomic DNA of barley has been evaluated by re-sequencing BAC clones with GS20 technology producing reads with an average 100 nucleotides (Wicker et al. 2006 and Fig. 24.1). The results showed that BAC inserts were covered very evenly by sequence and, at an overall coverage of 5–10-fold, every region of the BAC sequence was hit by at least a single read. At substantial oversampling (sequence coverage in the range of 60-fold), the sequences assembled into contigs of less than 100 bp up to more than 20 Kb (Roche-454 sequence assembly tool Newbler) which covered accurately the genic regions of the BAC clones. Subsequently, it has been shown that shotgun reads representing about 20-fold coverage generated for BAC clones on the

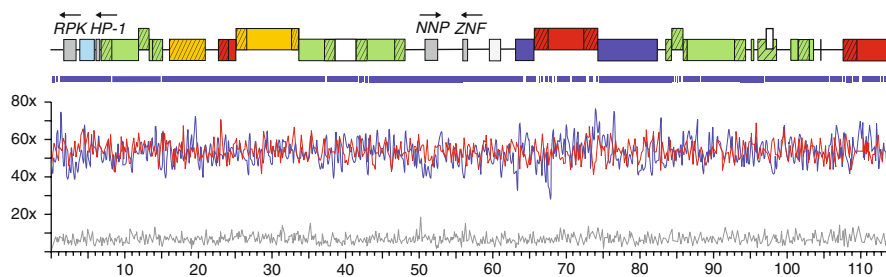


Fig. 24.1 Comparison of 454 GS20 and Sanger sequencing of barley BAC clones. A barley BAC clone previously Sanger sequenced (Genbank ID AF474072) was resequenced by 454 GS20 technology. Reads were assembled *de novo* by the Roche/454 assembly software Newbler. The upper panel indicates the annotation of genes (grey boxes) and repetitive elements (colored boxes) of the assembled contiguous Sanger sequence. Transposable elements that inserted into existing sequences are raised above their targets. Solid blue lines below the Sanger sequence annotation panel indicate regions represented by assembled GS20 read contigs. The BAC clone was sequence to 52.7x coverage and the distribution of reads (lower panel) is shown for the obtained reads (blue line) and for a simulated amount of reads (red line) in order to indicate the very even representation of the entire sequence by individual reads. Even at lower ~5x coverage (black line, lower panel) the entire BAC insert is covered by reads. (Figure taken from Wicker et al. 2006) (See Color Insert)

upgraded Roche-454 platform (GSFLX) assemble into few large contigs – in some cases into a single contig – representing the entire BAC insert (Stein et al. unpublished results). If individual clones are tagged specifically after random shearing by barcode adaptors (Binladen et al. 2007; Meyer et al. 2008; Meyer et al. 2007; Parameswaran et al. 2007), 50 BAC clones with an average 100 kb insert size could be pooled and sequenced in a single GSFLX run to obtain a sequence with an average 20-fold coverage per individual BAC clone. *De novo* 454 shotgun sequencing and sequence assembly of Triticeae genomic sequences (BAC clones, individual or Minimal Tiling Paths from BAC contigs) will become an important component of any strategy for sequencing the complex Triticeae genomes. It may even become routine and could potentially replace Sanger sequencing. This looks even more likely with the new Titanium system that promises to provide >400 nt read length and 0.5 Gb run capacity.

24.5 The Future of Triticeae Genome Sequencing

Today, we have gained already substantial insight into a large portion of the expressed genes of wheat and barley, into genome organization and, to some extent, into gene distribution patterns of parts of the Triticeae genomes. But, we are still far from having a complete picture of the genomes of our most important crop species genomes. These will be a major prerequisite before the full potential of these important crop species can be exploited to cope with future challenges in agriculture. The ultimate goal, thus, is to gain access to all genes and regulatory elements of wheat and barley in the context of a highly accurate reference genome sequence. The repertoire of alternative strategies to reach this goal has been reinforced significantly with the possibilities offered by NGS technologies.

Current NGS technologies provide new opportunities for targeted sequencing of the expressed portion of genomes. Hundreds of thousands of ESTs can be accumulated today for a few cents per individual read – a fraction of the cost less than a decade ago by Sanger EST sequencing (Cheung et al. 2006; Emrich et al. 2007; Weber et al. 2007). Comparisons of cDNA sequences produced from multiple sources, tissues, and from different genotypes rapidly will provide a rich source of information for the survey of gene-based sequence polymorphisms and identification of informative markers for breeding (Barbazuk et al. 2007; Brockman et al. 2008).

EST sequencing, however, can provide only limited information due to the underrepresentation of 5′ untranslated and untranscribed regions of genes as well as the lack of access to gene structure information. High-throughput sequencing combined with so-called sequence capture of selected genic regions may help to overcome this bottleneck. It has been demonstrated independently that specific target sequences (i.e. exons of thousands of genes or several Mb of tiled genomic sequence) can be captured efficiently by hybridization of sheared

genomic DNA to custom made oligo microarrays. Such DNA can be utilized as a template for high-throughput sequencing (Albert et al. 2007; Hodges et al. 2007; Okou et al. 2007). Thus, oligonucleotide tiling arrays representing the current unigene sets of wheat or barley can be used already to capture unprecedented sets of gene structure information for the Triticeae without the need for whole genome sequencing.

Since genes occupy only about 5% of the overall Triticeae genomes, is it justifiable to decipher the entire genomes including the 95% that is almost exclusively composed of repetitive DNA? Similar questions have accompanied almost every genome project. Repetitive DNA has been termed “junk DNA” (Ohno 1972) since its impact to species biology and phenotypes was largely unknown, raising the possibility that the majority of it could be dispensable. Is it, thus, dispensable in the context of Triticeae genome sequencing?

Although sequencing only the “gene space” sounds like a straightforward and less expensive approach, it would provide only a preliminary and fragmented mosaic of the whole picture. Non-genic regions consist mainly of intact or fragmented mobile elements (transposons, retrotransposons) present at many copies per genome. Mobile elements (repetitive DNA) have been identified as integral components of genes in mammals (Britten 2004; Cordaux et al. 2006) as well as in plants (Wang et al. 2006; Xiao et al. 2008). They contribute a significant level of polymorphism in eukaryotic genomes due to their mode of action (transposition, multiplication, and as targets for recombination) as has been observed in maize, rice, and barley (overview in Morgante 2006) and regulatory sequences of genes may be located in conserved non-coding regions far distant from the actual gene responsible for the trait (Clark et al. 2006; Salvi et al. 2007). Mobile elements constitute an important evolutionary force and, similar to genes, they can be under selective pressure (for review: Biémont and Vieira 2006; Kondrashov 2005; Morgante 2006). Consequently, if we focus only on the genes in Triticeae genome sequencing, we will not gain access in the long-term to all of the information necessary to obtain full understanding of wheat and barley performance as crops.

We, clearly, see the need for access to high-quality whole genome reference sequences of the Triticeae genomes. The most efficient and perhaps only feasible strategy to approach such an undertaking has been exemplified by international consortia such as the Human Genome Sequencing Consortium (<http://www.genome.gov/11006939>) and the International Rice Genome Sequencing Project (<http://rgp.dna.affrc.go.jp/IRGSP/>). Similarly, in the field of Triticeae structural genomics, the process of joining forces triggered the formation of the international genome consortia for wheat (IWGSC, <http://www.wheatgenome.org>) and barley (IBSC, <http://barleygenome.org>). Both scientific communities concluded that sequencing the Triticeae genomes in a clone-by-clone-based sequencing strategy supported by a genetically anchored physical map provides the most conservative but most appropriate step towards a “gold standard” reference sequence for these genomes. The development of physical maps of the whole genome (barley, *Aegilops tauschii*) or for all chromosomes individually

(*Triticum aestivum*) has been initiated and addressed (see Chapter 11). High density genetic maps for anchoring these physical maps are under development (see Chapters 7 and 11). The new possibilities offered through NGS technology are under constant evaluation (i.e. Wicker et al. 2008; Wicker et al. 2006) and likely will become an important component of whole Triticeae genome sequencing. Advances in sequencing technology will be carefully considered for any necessary revision of the plan to sequence wheat and barley. Thus, the wheat and barley research communities prepare to reach for their common goal of determining the genomic sequences of these important crop species.

24.6 Outlook

Innovations in sequencing technology induced a revolution in the way we consider approaches to molecular genetic analyses for any kind of organism. The full spectrum of technical applications that become available with high-throughput sequencing technologies can be compared to the impact of Polymerase Chain Reaction (PCR) soon after its introduction in the 1980s. Thus, it is likely that within the next five years some wheat chromosomes and possibly the barley genome will be sequenced. What will we do with this information?

Access to a high quality reference sequence of the wheat and barley genomes would enable whole genome based studies. Re-sequencing of diverse genotypes/germplasm revealed millions of single nucleotide polymorphisms in humans (The International HapMap Consortium 2007) and *Arabidopsis* (Clark et al. 2007; Ossowski et al. 2008) that can be exploited for genome-wide studies of linkage disequilibrium/association genetics in order to elucidate the basis of complex traits (see also Chapter 22).

Resequencing of the human genome and comparison to the high quality reference sequence revealed that so-called “copy-number-variations” (CNV) are a very frequent feature contributing a significant amount of overall polymorphism to the human pan-genome (Jakobsson et al. 2008). The extent to which this type of polymorphism is also characteristic of the wheat and barley genomes and the role that such genome variations may play in the expression of important agronomic traits is largely unknown. We know, already, that for the grapevine, large families related to wine characteristics have a higher copy number than in other plants (The French–Italian Public Consortium for Grapevine Genome Characterization 2007).

Although sequencing of the wheat and barley genomes are the major tasks to be tackled in the near future, major advancements and innovations will be needed in parallel to make the most efficient use of the genomes. With new systems for managing and exploiting Triticeae genetic resources, improving our capabilities and capacities in high-throughput precision phenotyping, and developing predictive models of gene regulatory networks and whole plant biology, there are exciting prospects ahead for improving wheat and barley.

References

- Albert, T.J., Molla, M.N., Muzny, D.M., Nazareth, L., Wheeler, D., Song, X., Richmond, T.A., Middle, C.M., Rodesch, M.J., Packard, C.J. et al. (2007) Direct selection of human genomic loci by microarray hybridization. *Nat. Meth.* 4, 903.
- Anderson, O.D., Rausch, C., Moullet, O. and Lagudah, E.S. (2003) The wheat D-genome HMW-glutenin locus: BAC sequencing gene distribution, and retrotransposon clusters. *Funct. Integr. Genomics* 3, 56–68.
- Barbazuk, W.B., Emrich, S.J., Chen, H.D., Li, L. and Schnable, P.S. (2007) SNP discovery via 454 transcriptome sequencing. *Plant J.* 51, 910–918.
- Barry, G.F. (2001) The use of the Monsanto draft rice genome sequence in research. *Plant Physiol.* 125, 1164–1165.
- Bennett, S.T., Barnes, C., Cox, A., Davies, L. and Brown, C. (2005) Toward the \$1000 human genome. *Pharmacogenomics* 6, 373–382.
- Bhat, P.R., Lukaszewski, A., Cui, X., Xu, J., Svensson, J.T., Wanamaker, S., Waines, J.G. and Close, T.J. (2007) Mapping translocation breakpoints using a wheat microarray. *Nucl. Acids Res.* 35, 2936–2943.
- Biémont, C. and Vieira, C. (2006) Junk DNA as an evolutionary force. *Nature* 443, 521–524.
- Binladen, J., Gilbert, M.T.P., Bollback, J.P., Panitz, F., Bendixen, C., Nielsen, R. and Willerslev, E. (2007) The use of coded PCR primers enables high-throughput sequencing of multiple homolog amplification products by 454 parallel sequencing. *PLoS ONE* 2, e197.
- Britten, R.J. (2004) Coding sequences of functioning human genes derived entirely from mobile element sequences. *Proc Natl Acad Sci USA* 101, 16825–16830.
- Brockman, W., Alvarez, P., Young, S., Garber, M., Giannoukos, G., Lee, W.L., Russ, C., Lander, E.S., Nusbaum, C. and Jaffe, D.B. (2008) Quality scores and SNP detection in sequencing-by-synthesis systems. *Genome Res.* gr.070227.070107.
- Buell, C.R. (2002) Obtaining the sequence of the rice genome and lessons learned along the way. *Trends Plant Sci.* 7, 538–542.
- Burr, B. (1999) A report from Singapore, September 1997: An international collaboration to sequence the rice genome. *Oryza: Newsletter for International Rice Genome Sequencing Project* 1999, 4–9.
- Burr, B. (2002) Mapping and sequencing the rice genome. *Plant Cell* 14, 521–523.
- Chaisson, M.J. and Pevzner, P.A. (2008) Short read fragment assembly of bacterial genomes. *Genome Res.* 18, 324–330.
- Cheung, F., Haas, B.J., Goldberg, S.M., May, G.D., Xiao, Y. and Town, C.D. (2006) Sequencing *Medicago truncatula* expressed sequenced tags using 454 Life Sciences technology. *BMC Genomics* 7, 272.
- Clark, R.M., Schweikert, G., Toomajian, C., Ossowski, S., Zeller, G., Shinn, P., Warthmann, N., Hu, T.T., Fu, G., Hinds, D.A. et al. (2007) Common sequence polymorphisms shaping genetic diversity in *Arabidopsis thaliana*. *Science* 317, 338–342.
- Clark, R.M., Wagler, T.N., Quijada, P. and Doebley, J. (2006) A distant upstream enhancer at the maize domestication gene *tb1* has pleiotropic effects on plant and inflorescent architecture. *Nat. Genet.* 38, 594–597.
- Close, T.J., Wanamaker, S.I., Caldo, R.A., Turner, S.M., Ashlock, D.A., Dickerson, J.A., Wing, R.A., Muehlbauer, G.J., Kleinhofs, A. and Wise, R.P. (2004) A new resource for cereal genomics: 22 K barley GeneChip comes of age. *Plant Physiol.* 134, 960–968.
- Cokus, S.J., Feng, S., Zhang, X., Chen, Z., Merriman, B., Haudenschild, C.D., Pradhan, S., Nelson, S.F., Pellegrini, M. and Jacobsen, S.E. (2008) Shotgun bisulphite sequencing of the *Arabidopsis* genome reveals DNA methylation patterning. *Nature* 452, 215–219.
- Cordaux, R., Udit, S., Batzer, M.A. and Feschotte, C. (2006) Birth of a chimeric primate gene by capture of the transposase gene from a mobile element. *Proc Natl Acad Sci USA* 103, 8101–8106.

- Devos, K.M., Ma, J., Pontaroli, A.C., Pratt, L.H. and Bennetzen, J.L. (2005) Analysis and mapping of randomly chosen bacterial artificial chromosome clones from hexaploid bread wheat. *Proc Natl Acad Sci USA* 102, 19243–19248.
- Dolinski, K. and Botstein, D. (2005) Changing perspectives in yeast research nearly a decade after the genome sequence. *Genome Res.* 15, 1611–1619.
- Dubcovsky, J., Ramakrishna, W., SanMiguel, P.J., Busso, C.S., Yan, L.L., Shiloff, B.A. and Bennetzen, J.L. (2001) Comparative sequence analysis of colinear barley and rice bacterial artificial chromosomes. *Plant Physiol.* 125, 1342–1353.
- Emberton, J., Ma, J., Yuan, Y., SanMiguel, P. and Bennetzen, J.L. (2005) Gene enrichment in maize with hypomethylated partial restriction (HMPR) libraries. *Genome Res.* 15, 1441–1446.
- Emrich, S.J., Barbazuk, W.B., Li, L. and Schnable, P.S. (2007) Gene discovery and annotation using LCM-454 transcriptome sequencing. *Genome Res.* 17, 69–73.
- Gerstein, M.B., Bruce, C., Rozowsky, J.S., Zheng, D., Du, J., Korbel, J.O., Emanuelsson, O., Zhang, Z.D., Weissman, S. and Snyder, M. (2007) What is a gene, post-ENCODE? History and updated definition. *Genome Res.* 17, 669–681.
- Goff, S.A., Ricke, D., Lan, T.-H., Presting, G., Wang, R., Dunn, M., Glazebrook, J., Sessions, A., Oeller, P., Varma, H. et al. (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). *Science* 296, 92–100.
- Goffeau, A., Barrell, B.G., Bussey, H., Davis, R.W., Dujon, B., Feldmann, H., Galibert, F., Hoheisel, J.D., Jacq, C., Johnston, M. et al. (1996) Life with 6000 genes. *Science* 274, 546–567.
- Gronenborn, B. and Messing, J. (1978) Methylation of single-stranded DNA in vitro introduces new restriction endonuclease cleavage sites. *Nature* 272, 375–377.
- Hanada, K., Zhang, X., Borevitz, J.O., Li, W.-H. and Shiu, S.-H. (2007) A large number of novel coding small open reading frames in the intergenic regions of the *Arabidopsis thaliana* genome are transcribed and/or under purifying selection. *Genome Res.* 17, 632–640.
- Hardin, S., Gao, X., Briggs, J., Willson, R. and Tu, S.-C. (2004) Methods for real-time single molecule sequence determination. United States Patent, 7,329,492, Visigen Biotechnologies, Inc. (TX).
- Harris, T.D., Buzby, P.R., Babcock, H., Beer, E., Bowers, J., Braslavsky, I., Causey, M., Colonell, J., DiMeo, J., Efcavitch, J. W et al. (2008) Single-molecule DNA sequencing of a viral genome. *Science* 320, 106–109.
- Hernandez, D., Francois, P., Farinelli, L., Osteras, M. and Schrenzel, J. (2008) De novo bacterial genome sequencing: millions of very short reads assembled on a desktop computer. *Genome Res.* 10.1101/gr.072033.072107.
- Hodges, E., Xuan, Z., Balija, V., Kramer, M., Molla, M.N., Smith, S.W., Middle, C.M., Rodesch, M.J., Albert, T.J., Hannon, G.J. et al. (2007) Genome-wide in situ exon capture for selective resequencing. *Nat. Genet.* 39, 1522.
- Hutchison, C.A., III. (2007) DNA sequencing: bench to bedside and beyond. *Nucl. Acids Res.* 35, 6227–6237.
- International Rice Genome Sequencing Project. (2005) The map-based sequence of the rice genome. *Nature* 436, 793–800.
- Itoh, T., Tanaka, T., Barrero, R.A., Yamasaki, C., Fujii, Y., Hilton, P.B., Antonio, B.A., Aono, H., Apweiler, R., Bruskewich, R. et al. (2007) Curated genome annotation of *Oryza sativa* ssp. *japonica* and comparative genome analysis with *Arabidopsis thaliana*. *Genome Res.* 17, 175–183.
- Jakobsson, M., Scholz, S.W., Scheet, P., Gibbs, J.R., VanLiere, J.M., Fung, H.-C., Szpiech, Z.A., Degnan, J.H., Wang, K., Guerreiro, R. et al. (2008) Genotype, haplotype and copy-number variation in worldwide human populations. *Nature* 451, 998.
- Kelleher, C.T., Chiu, R., Bosdet, H.S.I.E., Krzywinski, M.I., Fjell, C.D., Wilkin, J., Yin, T., DiFazio, S.P., Ali, J., Asano, J.K. et al. (2007) A physical map of the highly heterozygous

- Populus* genome: integration with the genome sequence and genetic map and analysis of haplotype variation. *Plant J.* 50, 1063–1078.
- Kidd, J.M., Cooper, G.M., Donahue, W.F., Hayden, H.S., Sampas, N., Graves, T., Hansen, N., Teague, B., Alkan, C., Antonacci, F. et al. (2008) Mapping and sequencing of structural variation from eight human genomes. *Nature* 453, 56–64.
- Komatsuda, T., Pourkheirandish, M., He, C., Azhaguvél, P., Kanamori, H., Perovic, D., Stein, N., Graner, A., Wicker, T., Tagiri, A. et al. (2007) Six-rowed barley originated from a mutation in a homeodomain-leucine zipper I-class homeobox gene. *Proc Natl Acad Sci USA* 104, 1424–1429.
- Kondrashov, A.S. (2005) Fruitfly genome is not junk. *Nature* 437, 1106.
- Korlach, J., Marks, P.J., Cicero, R.L., Gray, J.J., Murphy, D.L., Roitman, D.B., Pham, T.T., Otto, G.A., Foquet, M. and Turner, S.W. (2008) Selective aluminum passivation for targeted immobilization of single DNA polymerase molecules in zero-mode waveguide nanostructures. *Proc Natl Acad Sci USA* 105, 1176–1181.
- Lamoureaux, D., Peterson, D.G., Li, W., Fellers, J.P. and Gill, B.S. (2005) The efficacy of Cot-based gene enrichment in wheat (*Triticum aestivum* L.). *Genome* 48, 1120–1126.
- Lazo, G.R., Chao, S., Hummel, D.D., Edwards, H., Crossman, C.C., Lui, N., Matthews, D.E., Carollo, V.L., Hane, D.L., You, F.M. et al. (2004) Development of an expressed sequence tag (EST) resource for wheat (*Triticum aestivum* L.): EST generation, unigene analysis, probe selection and bioinformatics for a 16,000-locus bin-delineated map. *Genetics* 168, 585–593.
- Mardis, E. (2008) The impact of next-generation sequencing technology on genetics. *Trends Genet.* 24, 133–141.
- Margulies, M., Egholm, M., Altman, W.E., Attiya, S., Bader, J.S., Bemben, L.A., Berka, J., Braverman, M.S., Chen, Y.-J., Chen, Z. et al. (2005) Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437, 376–380.
- Messing, J. (1999) The formation of an International Rice Genome Sequencing Project. *Oryza: Newsletter for International Rice Genome Sequencing Project*, 2–4.
- Messing, J., Bharti, A.K., Karlowski, W.M., Gundlach, H., Kim, H.R., Yu, Y., Wei, F., Fuks, G., Soderlund, C.A., Mayer, K.F.X. et al. (2004) Sequence composition and genome organization of maize. *Proc Natl Acad Sci USA* 101, 14349–14354.
- Messing, J., Gronenborn, B., Müller-Hill, B. and Hopschneider, H.P. (1977) Filamentous coliphage M13 as a cloning vehicle: insertion of a HindII fragment of the lac regulatory region in M13 replicative form in vitro. *Proc Natl Acad Sci USA* 74, 3642–3646.
- Messing, J. and Vieira, J. (1982) A new pair of M13 vectors for selecting either DNA strand of double-digest restriction fragments. *Gene* 19, 269–276.
- Metzker, M.L. (2005) Emerging technologies in DNA sequencing. *Genome Res.* 15, 1767–1776.
- Meyer, M., Stenzel, U. and Hofreiter, M. (2008) Parallel tagged sequencing on the 454 platform. *Nat. Protocols* 3, 267–278.
- Meyer, M., Stenzel, U., Myles, S., Prufer, K. and Hofreiter, M. (2007) Targeted high-throughput sequencing of tagged nucleic acid samples. *Nucl. Acids Res.* 35, e97.
- Ming, R., Hou, S., Feng, Y., Yu, Q., Dionne-Laporte, A., Saw, J.H., Senin, P., Wang, W., Ly, B.V., Lewis, K.L.T. et al. (2008) The draft genome of the transgenic tropical fruit tree papaya (*Carica papaya* Linnaeus). *Nature* 452, 991–996.
- Morgante, M. (2006) Plant genome organisation and diversity: the year of the junk! *Curr. Opin. Biotechnol.* 17, 168–173.
- Ogihara, Y., Mochida, K., Nemoto, Y., Murai, K., Yamazaki, Y., Shin-I, T. and Kohara, Y. (2003) Correlated clustering and virtual display of gene expression patterns in the wheat life cycle by large-scale statistical analyses of expressed sequence tags. *Plant J.* 33, 1001–1011.
- Ohno, S. (1972) So much 'junk' DNA in our genome. *Brookhaven Symp. Biol.* 23, 366–370.
- Okou, D.T., Steinberg, K.M., Middle, C., Cutler, D.J., Albert, T.J. and Zwick, M.E. (2007) Microarray-based genomic selection for high-throughput resequencing. *Nat. Meth.* 4, 907.

- Olson, M.V. (2008) Human genetics: Dr Watson's base pairs. *Nature* 452, 819–820.
- Ossowski, S., Schneeberger, K., Clark, R.M., Lanz, C., Warthmann, N. and Weigel, D. (2008) Sequencing of natural strains of *Arabidopsis thaliana* with short reads. *Genome Res.* gr. 10.1101/080200.080108.
- Palmer, L. and McCombie, W.R. (2002) On the importance of being finished. *Genome Biol.* 3, comment2010.2011–comment2010.2014.
- Parameswaran, P., Jalili, R., Tao, L., Shokralla, S., Gharizadeh, B., Ronaghi, M. and Fire, A.Z. (2007) A pyrosequencing-tailored nucleotide barcode design unveils opportunities for large-scale sample multiplexing. *Nucl. Acids Res.* 35, e130.
- Paux, E., Roger, D., Badaeva, E., Gay, G., Bernard, M., Sourdille, P. and Feuillet, C. (2006) Characterizing the composition and evolution of homoeologous genomes in hexaploid wheat through BAC-end sequencing on chromosome 3B. *Plant J.* 48, 463–474.
- Peterson, D.G., Schulze, S.R., Sciara, E.B., Lee, S.A., Bowers, J.E., Nagel, A., Jiang, N., Tibbitts, D.C., Wessler, S.R. and Paterson, A.H. (2002) Integration of Cot analysis, DNA cloning, and high-throughput sequencing facilitates genome characterization and gene discovery. *Genome Res.* 12, 795–807.
- Pheasant, M. and Mattick, J.S. (2007) Raising the estimate of functional human sequences. *Genome Res.* 17, 1245–1253.
- Pop, M. and Salzberg, S.L. (2008) Bioinformatics challenges of new sequencing technology. *Trends Genet.* 24, 142–149.
- Qi, L.L., Echalié, B., Chao, S., Lazo, G.R., Butler, G.E., Anderson, O.D., Akhunov, E.D., Dvorak, J., Linkiewicz, A.M., Ratnasiri, A. et al. (2004) A chromosome bin map of 16,000 expressed sequence tag loci and distribution of genes among the three genomes of polyploid wheat. *Genetics* 168, 701–712.
- Rabinowicz, P.D. (2003) Constructing gene-enriched plant genomic libraries using methylation filtration technology. *Methods Mol. Biol.* 236, 21–36.
- Rabinowicz, P.D. and Bennetzen, J.L. (2006) The maize genome as a model for efficient sequence analysis of large plant genomes. *Curr. Opin. Plant Biol.* 9, 149–156.
- Rabinowicz, P.D., Citek, R., Budiman, M.A., Nunberg, A., Bedell, J.A., Lakey, N., O'Shaughnessy, A.L., Nascimento, L.U., McCombie, W.R. and Martienssen, R.A. (2005) Differential methylation of genes and repeats in land plants. *Genome Res.* 15, 1431–1440.
- Ramakrishna, W., Dubcovsky, J., Park, Y.J., Busso, C., Emberton, J., SanMiguel, P. and Bennetzen, J.L. (2002) Different types and rates of genome evolution detected by comparative sequence analysis of orthologous segments from four cereal genomes. *Genetics* 162, 1389–1400.
- Ronaghi, M. (2001) Pyrosequencing sheds light on DNA sequencing. *Genome Res.* 11, 3–11.
- Rostoks, N., Mudie, S., Cardle, L., Russell, J., Ramsay, L., Booth, A., Svensson, J., Wanamaker, S., Walia, H., Rodriguez, E. et al. (2005) Genome-wide SNP discovery and linkage analysis in barley based on genes responsive to abiotic stress. *Mol. Gen. Genomics* 274, 515–527.
- Rostoks, N., Park, Y., Ramakrishna, W., Ma, J., Druka, A., Shiloff, B., SanMiguel, P., Jiang, Z., Bruggeman, R., Sandhu, D. et al. (2002) Genomic sequencing reveals gene content, genomic organization, and recombination relationships in barley. *Funct. Integr. Genomics* 2, 51–59.
- Rudd, S. (2003) Expressed sequence tags: alternative or complement to whole genome sequences? *Trends Plant Sci.* 8, 321.
- Ryan, D., Rahimi, M., Lund, J., Mehta, R. and Parviz, B. (2007) Toward nanoscale genome sequencing. *Trends Biotechnol.* 25, 385–389.
- Salvi, S., Sponza, G., Morgante, M., Tomes, D., Niu, X., Fengler, K.A., Meeley, R., Ananiev, E.V., Svtashev, S., Bruggemann, E. et al. (2007) Conserved noncoding genomic sequences associated with a flowering-time quantitative trait locus in maize. *Proc. Natl. Acad. Sci. USA* 104, 11376–11381.

- Sanger, F. (1988) Sequences, sequences, and sequences. *Annu. Rev. Biochem.* 57, 1–29.
- Sanger, F. (2001) The early days of DNA sequences. *Nat. Med.* 7, 267–268.
- Sanger, F., Nicklen, S. and Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74, 5463–5467.
- Sasaki, T. and Burr, B. (1999) Rice genome sequencing as a gold mining for all. *Oryza: Newsletter for International Rice Genome Sequencing Project*, 1–2.
- Sasaki, T. and Burr, B. (2000) International Rice Genome Sequencing Project: the effort to completely sequence the rice genome. *Curr. Opin. Plant Biol.* 3, 138–142.
- Sasaki, T., Matsumoto, T., Yamamoto, K., Sakata, K., Baba, T., Katayose, Y., Wu, J., Niimura, Y., Cheng, Z., Nagamura, Y. et al. (2002) The genome sequence and structure of rice chromosome 1. *Nature* 420, 312–316.
- Service, R.F. (2006) Gene sequencing: the race for the \$1000 genome. *Science* 311, 1544–1546.
- Shendure, J. and Ji, H. (2008) Next-generation DNA sequencing. *Nat. Biotechnol.* 26, 1135–1145.
- Shendure, J., Mitra, R.D., Varma, C. and Church, G.M. (2004) Advanced sequencing technologies: methods and goals. *Nat. Rev. Genet.* 5, 335–343.
- Somerville, C. and Koornneef, M. (2002) A fortunate choice: the history of *Arabidopsis* as a model plant. *Nat. Rev. Genet.* 3, 883–889.
- Stein, N. (2007) Triticeae genomics: advances in sequence analysis of large genome cereal crops. *Chromosome Res.* 15, 21–31.
- Stein, N., Prasad, M., Scholz, U., Thiel, T., Zhang, H., Wolf, M., Kota, R., Varshney, R.K., Perovic, D. and Graner, A. (2007) A 1000 loci transcript map of the barley genome – new anchoring points for integrative grass genomics. *Theor. Appl. Genet.* 114, 823–839.
- Swarbreck, D., Wilks, C., Lamesch, P., Berardini, T.Z., Garcia-Hernandez, M., Foerster, H., Li, D., Meyer, T., Muller, R., Ploetz, L. et al. (2008) The *Arabidopsis* Information Resource (TAIR): gene structure and function annotation. *Nucl. Acids Res.* 36, D1009–D1014.
- The *C. elegans* Sequencing Consortium. (1998) Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science* 282, 2012–2018.
- The ENCODE Project Consortium. (2007) Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* 447, 799–816.
- The French–Italian Public Consortium for Grapevine Genome Characterization. (2007) The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nature* 449, 463–467.
- The International HapMap Consortium. (2007) A second generation human haplotype map of over 3.1 million SNPs. *Nature* 449, 851–861.
- Thurman, R.E., Day, N., Noble, W.S. and Stamatoyannopoulos, J.A. (2007) Identification of higher-order functional domains in the human ENCODE regions. *Genome Res.* 17, 917–927.
- Timko, M., Rushton, P., Laudeman, T., Bokowiec, M., Chipumuro, E., Cheung, F., Town, C. and Chen, X. (2008) Sequencing and analysis of the gene-rich space of cowpea. *BMC Genomics* 9, 103.
- Travis, J. (2008) Uncorking the grape genome. *Science* 320, 475–477.
- Tuskan, G.A., DiFazio, S., Jansson, S., Bohlmann, J., Grigoriev, I., Hellsten, U., Putnam, N., Ralph, S., Rombauts, S., Salamov, A. et al. (2006) The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* 313, 1596–1604.
- Velasco, R., Zharkikh, A., Troggio, M., Cartwright, D.A., Cestaro, A., Pruss, D., Pindo, M., FitzGerald, L.M., Vezzulli, S., Reid, J. et al. (2007) A high quality draft consensus sequence of the genome of a heterozygous grapevine variety. *PLoS ONE* 2, e1326.
- Venter, J.C., Smith, H.O. and Hood, L. (1996) A new strategy for genome sequencing. *Nature* 381, 364.
- Vij, S., Gupta, V., Kumar, D., Vydianathan, R., Raghuvanshi, S., Khurana, P., Khurana, J.P. and Tyagi, A.K. (2006) Decoding the rice genome. *BioEssays* 28, 421–432.

- Wang, Y., Tang, X., Cheng, Z., Mueller, L., Giovannoni, J. and Tanksley, S.D. (2006) Euchromatin and pericentromeric heterochromatin: comparative composition in the tomato genome. *Genetics* 172, 2529–2540.
- Warren, W.C., Hillier, L.W., Graves, J.A.M., Birney, E., Ponting, C.P., Grützner, F., Belov, K., Miller, W., Clarke, L., Chinwalla, A.T. et al. (2008) Genome analysis of the platypus reveals unique signatures of evolution. *Nature* 453, 175–183.
- Weber, A.P.M., Weber, K.L., Carr, K., Wilkerson, C. and Ohlrogge, J.B. (2007) Sampling the *Arabidopsis* transcriptome with massively parallel pyrosequencing. *Plant Physiol.* 144, 32–42.
- Wei, F., Coe, E., Nelson, W., Bharti, A.K., Engler, F., Butler, F., Kim, H.R., Goicoechea, J.L., Chen, M., Lee, S. et al. (2008) Physical and genetic structure of the maize genome reflects its complex evolutionary history. *PLoS Genet* 3, e123.
- Wheeler, D.A., Srinivasan, M., Egholm, M., Shen, Y., Chen, L., McGuire, A., He, W., Chen, Y.-J., Makhijani, V., Roth, G.T. et al. (2008) The complete genome of an individual by massively parallel DNA sequencing. *Nature* 452, 872–876.
- Wicker, T., Matthews, D. and Keller, B. (2002) TREP, a database for Triticeae repetitive elements. *Trends Plant Sci.* 7, 561–562.
- Wicker, T., Narechania, A., Sabot, F., Stein, J., Vu, G.T.H., Graner, A., Ware, D. and Stein, N. (2008) Low-pass shotgun sequencing of the barley genome facilitates rapid identification of genes, conserved non-coding sequences and novel repeats. *BMC Genomics* 9, 518.
- Wicker, T., Schlagenhauf, E., Graner, A., Close, T., Keller, B. and Stein, N. (2006) 454 sequencing put to the test using the complex genome of barley. *BMC Genomics* 7, 275.
- Wicker, T., Stein, N., Albar, L., Feuillet, C., Schlagenhauf, E. and Keller, B. (2001) Analysis of a contiguous 211 kb sequence in diploid wheat (*Triticum monococcum* L.) reveals multiple mechanisms of genome evolution. *Plant J.* 26, 307–316.
- Wicker, T., Yahiaoui, N., Guyot, R., Schlagenhauf, E., Liu, Z.D., Dubcovsky, J. and Keller, B. (2003) Rapid genome divergence at orthologous low molecular weight glutenin loci of the A and A(m) genomes of wheat. *Plant Cell* 15, 1186–1197.
- Wicker, T., Zimmermann, W., Perovic, D., Paterson, A.H., Ganai, M., Graner, A. and Stein, N. (2005) A detailed look at 7 million years of genome evolution in a 439 kb contiguous sequence at the barley *Hv-eIF4E* locus: recombination, re-arrangements and repeats. *Plant J.* 41, 184–194.
- Wilhelm, B.T., Marguerat, S., Watt, S., Schubert, F., Wood, V., Goodhead, I., Penkett, C.J., Rogers, J. and Bahler, J. (2008) Dynamic repertoire of a eukaryotic transcriptome surveyed at single-nucleotide resolution. *Nature* 453, 1239–1243.
- Wilson, R.K., Ware, D., Wing, R.A., McCombie, W.R., Schnable, P.S., Clifton, S.W., Aluru, S., Stein, L.D., Martienssen, R. and Fulton, R. (2007) Sequencing the maize B73 genome. Progress Report (www.maizgedb.org/sequencing_project.php).
- Wu, R. and Kaiser, A.D. (1968) Structure and base sequence in the cohesive ends of bacteriophage lambda DNA. *J. Mol. Biol.* 35, 523–527.
- Wu, R. and Taylor, E. (1971) Nucleotide sequence analysis of DNA. II. Complete nucleotide sequence of the cohesive ends of bacteriophage lambda DNA. *J. Mol. Biol.* 57, 491–511.
- Xiao, H., Jiang, N., Schaffner, E., Stockinger, E.J. and van der Knaap, E. (2008) A retro-transposon-mediated gene duplication underlies morphological variation of tomato fruit. *Science* 319, 1527–1530.
- Yamada, K., Lim, J., Dale, J.M., Chen, H., Shinn, P., Palm, C.J., Southwick, A.M., Wu, H.C., Kim, C., Nguyen, M. et al. (2003) Empirical analysis of transcriptional activity in the *Arabidopsis* genome. *Science* 302, 842–846.
- Yan, L., Loukoianov, A., Tranquilli, G., Helguera, M., Fahima, T. and Dubcovsky, J. (2003) Positional cloning of the wheat vernalization gene *VRN1*. *Proc. Natl. Acad. Sci. USA* 100, 6263–6268.

- Yu, J., Hu, S., Wang, J., Wong, G.K.-S., Li, S., Liu, B., Deng, Y., Dai, L., Zhou, Y., Zhang, X. et al. (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). *Science* 296, 79–92.
- Yu, J., Wang, J., Lin, W., Li, S., Li, H., Zhou, J., Ni, P., Dong, W., Hu, S., Zeng, C. et al. (2005) The genomes of *Oryza sativa*: a history of duplications. *PLoS Biol* 3, e38.
- Zhang, H., Sreenivasulu, N., Weschke, W., Stein, N., Rudd, S., Radchuk, V., Potokina, E., Scholz, U., Schweizer, P., Zierold, U. et al. (2004) Large-scale analysis of the barley transcriptome based on expressed sequence tags. *Plant J.* 40, 276–290.

Index

A

- Abscisic acid (ABA), 517
- Advanced-backcross QTL approach (ABQA), 518
- Aegilops tauschii* (goat grass), 35–36, 90
 - core collection, 61
 - databases, 61
 - institutions with accessions, 60
- × *Aegilotriticum*, 36
- Agrobacterium*-mediated transformation
 - methods, 372–373, 437–438
- Agropyron cristatum*, 36
- Agropyron desertorum*, 36
- Agropyron fragile*, 36
- Agropyron michnoi*, 36
- Agropyron mongolicum*, 36
- Agropyron* (wheatgrass), 36–37
 - accessions per species, 67
 - databases, 67
 - institutions with accessions, 67
- All-Russian Scientific Research Institute of Plant Industry (VIR), 82
- Al tolerance, 527
- Aluminium (Al) toxicity, 526
- Amblyopyrum muticum*, 37
- Amplified Fragment Length Polymorphisms (AFLPs), 145, 205, 261, 344
 - fingerprinting, 261
 - markers, 221
 - steps, 260–261
- Aneuploidy, 121
- Annotation methods, 407
 - ancillary, 408
 - automated and semi-automated, 413
 - comparative genome, 413–414
 - manual, 412–413
- Antisense ODNs, 369–370

- Arabidopsis thaliana*, 720
- Arabidopsis* transformation, 435–436
- Association mapping studies, 482
 - high-resolution, 512

B

- BAC contigs, 347–348
- BAC landing, 441
- Bacterial Artificial Chromosome (BAC)
 - libraries, 269–272, 330, 331, 345, 440
 - characteristics of, 270–271
 - chromosome arm-specific, 305
 - physical map construction, 304
 - positional cloning, 305
 - rye, 272
 - wheat and barley, 269
- Barley Genes and Genetic Stocks (BGS), 57
- BARley REtrotransposons (BARE), 460
- Barley Stripe Mosaic Virus (BSMV), 351, 366, 568
- Barley yellow dwarf virus* (BYDV), 560, 573
 - economic importance, 578–579
 - high quality sequences, 578–579
 - multiple sequencing, 578
- Basic Local Alignment Search Tool (BLAST) search, 414
- Biotechnology, 202
- “Bioversity Directory of Germplasm Collections”, 33
- BLASTN search, 414
- BLASTX search, 414–415
- Blumeria graminis* (powdery mildew), 560
 - genome size, 577–578
 - hordei* (*Bgh*), 569

- Brachypodium distachyon*, 37, 203, 427
 accessions per taxon, 72
 applications as model for grass research
 as functional model, 444–446
 as structural model for wheat and
 barley genomics, 443–444
 comparison of models and crops, 429
 as experimental system
 characteristics of diploid inbred
 lines, 432
 chemical and radiation mutagenesis,
 435
 germplasm resources and natural
 diversity, 434–435
 growth requirements and flowering
 triggers, 432–434
 related species, 438–439
 transformation and T-DNA tagging,
 435–438
 future prospects and directions, 446
 gene function analyses, 444–446
 genome size and polyploidy, 430–431
 genomic resources
 BAC library resources, 440–441
 bioinformatic resources, 442
 cDNA libraries, 439–440
 ESTs, 439–440
 genome sequencing, 442
 physical and genetic maps, 441
 inbred lines, 431–433, 434–435, 437
 institutions with accessions, 72
 model for plant-pathogen
 interactions, 446
 model for plants-mycorrhizal fungi
 interactions, 446
 model for responses to wounding and
 insect attack, 446
 model for root diseases, 433
 model systems, 427–428, 444
 as new model genome, 349
 relationships and phenotypes, 429
 relationship to other grasses, 431
Brachypodium pinnatum, 37
 Bread wheat, 85, 89, 331
 genome of, 408
 homoeologous chromosomes, 85
 Brome mosaic virus (BMV), 368, 568
 Bulk segregant analysis (BSA), 221
- C**
 Candidate gene approach (CGA)
 shortcomings, 512
 CarteBlanche, 205
- Carthagine, 214
 Caryopsis, 94
 C-banding, 121, 123, 126, 128–129, 142
 idiogram representation, 128
 cDNA clones, 257
 cDNA libraries, 439–440
 Cereal genomics, 255–256
 Cereals, 3, 81–82, 451–452
 phases of cytogenetic research, 121
 sub-families, 452
 Cereal \times *Triticosecale* (triticale), 31–32
 Chromosomal mutations, 146
 Chromosomal rearrangements, 257–258,
 457–458
- Chromosome
 analysis, 288
 BAC contig maps, 305
 deletion mapping, 146
 engineering, 127
 loss, 148
- Chromosome genomics
 advantages, 285
 applying flow cytogenetics, 290–301
 barley, 295–297
 hexaploid wheat, 291–294
 rye, 297–299
 tetraploid durum wheat, 294–295
 toolkit for Triticeae chromosome
 sorting, 299–301
 BAC contig physical maps and positional
 gene cloning, 302–305
 Bacterial Artificial Chromosome (BAC)
 libraries, 301–302
 cytogenetic mapping and chromosome
 structure, 308–309
 development of molecular markers,
 305–307
 flow cytogenetics, 288–289
 principles of flow cytometry and
 sorting, 289
 molecular organization of subgenomic
 regions, 305
 physical and genetic mapping using flow-
 sorted chromosomes, 307–308
 relative and molecular sizes, 286
 Chromosome-specific markers, 142
 Chromosome walking process, 327, 346
 Cluster analysis
 The Code, 7
 Codons Optimised to Discover Deleterious
 Lesions (CODDLE), 362
 Colinearity relationships, 456
 Community Sequencing Program, 442

- Comparative genomics, 452–453
 - application, 453
 - colinearity, 453
 - duplications, 464–467
 - genome evolution, 465
 - model for structural evolution, 466
 - rice and wheat, 465
 - segmental, 466
 - whole genome, 465
 - future prospectus, 470–471
 - genome scale
 - colinearity, 457
 - marker based macrocolinearity studies, 454–455
 - sequence based macrocolinearity studies, 455–458
 - locus-based level
 - genomic resources, 458
 - inter and intra specific microcolinearity studies, 459
 - interspecific comparative studies, 458–460
 - intraspecific comparisons, 460–462
 - intravarietal comparisons, 462–464
 - map-based cloning, 458
 - microcolinearity studies, 462
 - phylogenetic relationships between grasses and Triticeae species, 452
 - restriction fragment length polymorphisms (RFLP), 454
 - tool for gene discovery and marker development
 - colinearity-based gene cloning, 467–469
 - gene annotation and marker development, 469
 - increasing marker density, 469
 - map-based cloning, 468
 - reducing genetic interval, 468
 - Comparative mapping, 469
 - Composite interval mapping (CIM), 186
 - Consensus maps
 - challenge of optimising order of markers, 215
 - complications, 172
 - composite or framework map, 216
 - de novo, 172
 - issues in building, 214
 - Conserved Orthologous Set (COS) markers, 469
 - Consultative Group of International Agricultural Research (CGIAR), 33
 - Contig, 324
 - “Crop Strategies”, 33
 - Cross genome map-based cloning, 467
 - Crossover interference, 239
 - C-terminal domains, 625–626
 - Cytogenetics, 121–122
 - analyses, 137
 - Cytogenetics and genomics to hybridisations in *Hordeum*, 137–154
 - cytogenetics and species relationships, 142–145
 - FISH probes, 143
 - genomic composition, 145
 - interspecific genome affinities, 143–144
 - cytological characterisation and chromosome nomenclature, 138–142
 - chromosome sizes, 138
 - DNA probes, 140–141
 - genome, 139
 - geographical distribution, 138
 - generation of haploid barley, 147–149
 - chromosome loss, 148
 - interspecific crosses, 147
 - mitosis-dependent uniparental elimination, 148
 - physical mapping, 145–147
 - deletion-based, 146
 - integration into genetic maps, 146
 - in situ* hybridisation, 146–147
 - practical breeding applications, 149–154
 - addition lines, 149–150
 - frequency of homoeologous, 153
 - genomic *in situ* hybridisation, 152, 154
 - introgression lines, 151–154
 - numbers of ILs, 153
 - substitution lines, 150
 - Cytogenetic stock, 319–321
- D**
- DArT arrays, 265–266
 - Dasyphyrum villosum*, 37
 - Dehydration avoidance, 516
 - Dehydration tolerance, 516
 - Deletion-based physical mapping, 146
 - Deletion-bin mapping, 127
 - De novo* detection, 415
 - De novo* repeat identification software, 419

- Developmental and reproductive traits
 - advances in genomics and gene identification, 603–604
 - gene catalogues, 594
 - identifying flowering time genes
 - candidate gene method, 595
 - pathway regulating photoperiod response, 596
 - positional cloning/candidate gene hybrid method, 597–599
 - positional cloning method, 596–597
 - identifying inflorescence development genes
 - candidate gene method, 599
 - positional cloning method, 599–600
 - understanding gene function
 - analysis of genetic pathways, 600–602
 - models of vernalization pathway, 602
 - validation of candidate flowering genes, 602
 - using flowering and inflorescence genes in breeding, 605
 - variation in inflorescence architecture, 594
- Diversity Array Technology (DArT), 265, 307
- markers, 205, 207
- DNA sequencing, 388
- Domestication bottlenecks, 86–88
- Domestication in fertile crescent, 81–105
 - evolution and
 - barley, 90–92
 - rye, 92–94
 - wheat, 84–90
 - origins of cultivated plants and agriculture, 82–83
 - traits modified
 - brittle-rachis, 98
 - control of flowering time, 102
 - free-threshing, 94–97
 - grain hardness, 100–101
 - grass species, genome compositions and genes controlling, 95–96
 - heading time, 104
 - kernel rows in ear, 99
 - photoperiod, 102–103
 - plant height, 99–100
 - reduced seed dormancy, 102
 - seed size and grain yield, 98–99
 - tillering, 101
 - vernalization, 103–104
- Domestication syndrome, 83
- Dormancy, 102
- DotPlot, 416
- Doubled haploids, 147
- Double-stranded RNA (dsRNA), 369
 - interference, 565
- Drought
 - episode, 516
 - resistance, 516
 - stress, 516
 - tolerance
 - barley, 517–518
 - wheat, 519–521
- Durable resistance, 561
- E**
- EcoTILLING, 515
- Ectopic recombination, 692, 693
- Einkorn, 85–87
 - dispersed-specific model, 87
 - domestication bottleneck, 85–88
 - race β , 86
- Elymus canadensis*, 37
- Elymus caninus*, 37
- Elymus dahuricus*, 37
- Elymus elymoides*, 37
- Elymus fibrosus*, 38
- Elymus glaucus*, 38
- Elymus hoffmannii*, 38
- Elymus hystrix*, 38
- Elymus lanceolatus*, 38
- Elymus macrourus*, 38
- Elymus mutabilis*, 38
- Elymus nutans*, 38
- Elymus repens*, 38
- Elymus semicostatus*, 38
- Elymus sibiricus*, 38
- Elymus trachycaulus*, 38
- Elymus villosus*, 38
- Elymus virginicus*, 38
- Elymus* (Wheatgrass, Wild Rye), 37–38
 - accessions per taxon, 64–66
 - institutions with accessions, 64
- Embryogenic callus, 436
- EQTLs, 513
- Eremopyrum* (False Wheatgrass), 38
- Eremopyrum orientale*, 38
- Ergot of rye, 559
- Ethyl methane sulphonate (EMS), 360
 - mutagen, 435
- European Barley Database (EBDB), 56
- European Cereals Aneuploid Co-operative (EWAC), 52

- European Cooperative Programme for Plant Genetic Resources (ECPGR), 33, 50–51
- European Plant Genetic Resources Search Catalogue (EURISCO), 33
- European Research Area Networks - Plant Genomics (ERA-PG), 105
- European Rye Database of ECPGR, 63
- European Triticale Database (ETDB), 57
- European Wheat Database (EWDB), 51
- Evolutionary Optimization Strategy (ES), 164
- Expressed Sequenced Tags (ESTs), 256, 287, 409
 - application, 268
 - based markers, 222
 - and cDNA clones, 410
 - sequences, 268–269
 - SSR markers, 145, 259
- Expression QTL (eQTL) mapping, 564
- Ex situ* genebank collections, 31–32
 - accessions, 45
 - accessions per species, 69–71
 - Aegilops*, 59–61
 - Agropyron*, 66–67
 - Brachypodium*, 71–72
 - collections of genetic stocks and mutants, 46–49
 - countries and institutions, 44–46
 - data sources, 33
 - economic importance, 4
 - Elymus*, 64–66
 - genebanks, 47–48
 - genera and level of acceptance, 15–22
 - genera and species, 46
 - holdings by continents, 46
 - Hordeum*, 52–56
 - individual species, 51
 - institutions, 50, 68
 - other species, 67–71
 - Secale*, 62–63
 - × *Triticosecale*, 57–59
 - Triticum*, 49–52
 - core collections, 52
 - genetic stocks collections, 52
 - uses, 82
- F**
- FAO Institution Codes, 34
- Fast neutron radiation (FNR)
 - mutagen, 435
- Fertile Crescent agriculture, 83
 - origin of, 83
- FGENESH (gene finder), 411
- Fixed-effect model, 181
- Flow cytogenetic toolbox for Triticeae, 299
- Flow cytometry, 288
- Flowering time, 591–592
- Flow karyotype, 288
 - barley, 296
 - hexaploid bread wheat, 292
 - rye, 298
 - tetraploid durum wheat, 295
- Flow-sorted chromosomes and DNA applications, 301
 - DNA amplified, 306
- Fluorescence *in situ* hybridisation (FISH), 127, 139, 147, 318
- Food and Agriculture Organization (FAO), 32
- Founder package, 84
- Fractioned Pool Design (FPD), 189
- Framework-map approach, 216
- Free-threshing
 - evolution of, 97
 - mutations, 97
- Functional validation
 - stable genetic transformation
 - design of transformation vectors, 374–375
 - insertional mutagenesis, 376
 - linking manipulated gene expression with gene function, 376–378
 - patterns of DNA-integration, 373
 - transfer of recombinant DNA into plant cells, 371–373
 - Targeted Induced Local Lesions in Genomes (TILLING)
 - mutagens and mutation frequency, 360–362
 - mutation confirmation and functional validation, 365–366
 - mutation detection and validation, 364–365
 - mutation spectrum analysis, 362
 - populations for reverse genetics, 363–364
 - web-based computational tools, 362–363
 - transient gene validation assays
 - antisense oligodeoxynucleotide, 369–371
 - biolistic approaches, 368–369
 - virus induced gene silencing (VIGS), 366–368

Fusarium graminearum (*Fusarium* head blight or scab), 560, 574–575
 genomic sequence, 575
 genetic mapping, 575
 whole-genome shotgun, 574
Fusarium head blight (FHB) disease, 574

G

G×E interaction, 181
 Gene and repetitive sequence annotation
 genome sequence and annotation data
 comparative genome annotation, 413–414
 Poaceae species, 414
 Triticeae genomes, 410–411
 Triticeae transcriptome, 409–410
 repetitive sequences
 challenge of large number, 419–420
 methods for identification of
 transposable elements, 414–416
 problematic sequence in TEs, 418
 problems with transposable elements, 416–418
 software for repeat recognition and isolation, 418–419
 Triticeae genomics, 408–409
 complicating annotation process, 409
 Genebanks, 32
 data sources, 33
 genetic stocks collections, 46, 49
 Gene cloning, colinearity-based
 Brachypodium, 468–469
 cross genome map-based cloning, 467
 flanking genes, 468
 genome rearrangements, 468
 orthologous position, 467–468
 quantitative traits loci (QTL), 467
 Rht-1 gene, 467
 Gene-for-gene interactions, 573
 Gene-for-gene systems, 560
 Gene function
 heterologous transgene expression, 378
 ‘Gene island,’ 203–204
 Gene isolation, 467–468
 Gene model
 annotating, 412
 Generic binary vectors, 375
 Generic changes, 8–9
 Generic delimitation, 13
 Generic descriptions, 13
 Genetical genomics, 564
 Genetically Effective Cell Number (GECN), 363

Genetic analysis, methods for
 construction of high quality dense genetic maps
 building consensus maps, 171–174
 comparing quality characteristics, 172
 complication due to ‘pseudo-linkage’ and negative interference, 166–170
 increasing stability of multilocus maps, 170–171
 map verification procedures, 165–166
 multilocus mapping, 170
 multilocus ordering, 164–165
 negative interference, 169
 re-sampling based procedure, 166
 high-resolution mapping based on selective DNA pooling
 association analysis, 192–193
 example of association analysis, 194–195
 linkage analysis (RIL), 191–192
 parallel fitting of 2-QTL model, 194
 QTL mapping, 189–191
 RIL data analysis, 193–194
 simulated data for association analysis, 195
 simulations, 193
 QTL mapping
 approaches, 188
 detecting variance effect eQTL, 180
 joint analysis, 187
 multiple environments, 181–188
 multiple trait analysis, 175–178
 single interval vs. multiple interval mapping, 177
 single- vs. multiple-environment analysis, 185
 variance- covariance effect, 178–181
 Genetic architecture, 175
 Genetic distance, 317
 Genetic linkage maps, 202, 206–207
 high quality, 202
 with markers, 206–207
 micro-level comparative genomics sequence analyses, 203–204
 molecular markers, 204
 steps in construction of, 204
 use of, 201
 Genetic mapping, 201–224
 consensus maps, 213–216
 composite/framework map, 216
 order of markers, 215
 future directions, 223–224

- genetic linkage maps, 203–207
 - barley, 208–209
 - durum, 208
 - rye, 209
 - Triticale*, 209
 - wheat, 208
- high-resolution mapping, 219–223
 - purpose, 223
 - steps, 220
 - studies, 221
- map curation, 211–213
 - simulated chromosome, 213
- physical linkage maps, 210–211
 - HAPPY mapping, 211
 - optical mapping, 210
 - radiation induced deletion mutations, 210
- QTL mapping
 - molecular markers, 217
 - practical considerations for, 217–219
- Genetic maps, 145, 163, 201, 204, 223, 255, 256, 258, 259, 261, 265, 318, 324, 441
 - comparative mapping, 216
 - high-resolution mapping, 216, 219–223, 343
 - map-based cloning, 202, 216
 - marker-assisted selection, 216
 - merging, 172
 - quality, 201
 - factors affecting, 211–213
- Genetic marker, 343
- Genetic mutations, 363
- Genetic resources in *ex situ* genebank collections, 31–72
 - cultivated and useful species, 35–44
 - material and methods
 - handling of nomenclature, 34–35
 - information sources, 32–33
- Genetic Resources Information Network of USDA, 35
- Genetic stocks
 - collections, 49
 - and mutants, 46
- Genome annotation, 408
 - guidelines, 411
- Genome maps, 145
- Genome sequences, 442, 469, 514
 - initiatives, 408
- Genome structure and evolution
 - conservative and dynamic strata of Triticeae genomes, 698–699
 - gene order paradox, 697–698
- genome evolution
 - centromere-telomere axis of Triticeae chromosomes, 699–703
 - TEs and Triticeae genome evolution, 696–697
 - structure of Triticeae genome
 - genome size, 685–686
 - interspersed repeated sequences, 695–696
 - overall structure, 686–687
 - tandem repeated sequences, 687–692
- Genomic clones, 257
- Genomic *in situ* hybridization (GISH), 127
- Genomic rearrangements, 100
- Genomic resources
 - BAC libraries, 408
 - ESTs, 408
- Genomics approaches, 534–536
- Genomics- Assisted Analysis and Exploitation of Barley Diversity (EXBARDIV), 105
- Genomic sequence data, 407
- Genomics of biotic interactions
 - disease epidemics and current threats
 - integrative genomics, 561
 - plant defenses, 560–561
 - pathogen genomics
 - barley yellow dwarf virus (BYDV), 578–579
 - Blumeria graminis* (powdery mildew), 577–578
 - Fusarium graminearum* (*Fusarium* head blight), 574–575
 - Mycosphaerella graminicola* (*Septoria Tritici* blotch), 576
 - Puccinia graminis* (stem rust), 575–576
 - Stagonospora nodorum* (*Stagonospora Nodorum* blotch), 577
- toolbox for investigating
 - genes silenced or overexpressed, 567–568
 - high-throughput functional analysis, 565–568
 - high-throughput TIGS pipeline, 566
 - integration of phenotypic, genetic and physical-map data, 563–564
 - molecule profiling approaches, 562–563
- Triticeae-fungal host interactions, 569–571
 - role of R genes, 571

Genomics of biotic interactions (*cont.*)

Triticeae-fungal nonhost interactions,
571–573

Triticeae interactions with insects,
viruses, worms and bacteria,
573–574

Genomics of quality traits

barley quality

animal feed qualities, 612

β -amylase, 615–616

CesA and *Csl* gene, 613

characteristics assessing malt
quality, 615

germination as key variable, 619–621

germination behaviour, 620

human food, 612–614

malt extract QTL, 617

malting and brewing, 614–615

QTL associated with malting quality,
616–619

grain hardness and carbohydrates in
wheat and barley

grain hardness, 635–636

non-starch polysaccharides, 635

starch composition, 634

starch content, 633–634

traits not analysed at genomic level

grain protein content, 637–638

impact of new technologies, 638–639

milling yield, 636–637

water absorption, 637

wheat quality

assaying variation in seed storage
proteins, 629–630

finely divided bran specks in flour,
631–632

flour color, 630

flour paste viscosity, 633

flour proteins, 622–625

gluten polymer, 623

High Molecular Weight Glutenin
Subunits (HMWGS), 625–626

Low Molecular Weight Glutenin
Subunits (LMWGS), 626–627 seed
storage protein gene structure and
variation, 627–629

structure of wheat prolamin loci, 628
yellowness of flour and end products,
631

Genomics of tolerance to abiotic stress

genotype \times environment interaction,
534–535

microarray studies, 504–511

prospects of genomics-assisted
improvement, 535–537

QTLs and genes

aluminium toxicity, 526–528

boron toxicity, 528–529

drought, 516–521

high temperature, 533

low nutrients, 522–526

low temperature, 531–533

salinity, 521–522

tolerance to waterlogging, 530–531
zinc and manganese deficiency,
529–530

searching QTLs and genes

candidate gene approach, 512–513
exploiting -omics platforms,
513–515

studies on identification of genes,
496–503

studies on identification of QTLs,
483–495

Genomics of transposable elements

classification, 389

comparative, 399

exploitation as molecular markers,
399–400

functional, 395–398

direct effects, 395–398

effects on genes, sequence chimeras,
and gene regulation, 398

genome size and gene content, 389

structural, 390–395

DNA content ratio, 390

LINE target-primed reverse
transcription, 393

retrotransposition, 392

Rolling-Circle transposition
mechanism, 395

TE structures, 391

TIR Order transposition, 394

Genomics platforms, 481, 482

Genomic SSRs, 258

Genotype \times environment
interaction, 534

Germplasm resources

Giemsa C-banding patterns, 142

“Global Barley Register”, 54

Global Crop Diversity Trust, 33

Crop Strategies, 33, 46, 49

Wheat Strategy, 57, 62

“Global Crop Registers”, 49

Grain hardness/texture, 100

Green revolution, 99

GS1, 523
Gsp-1 genes, 101
 Guided Evolution Strategy, 165, 173

H

HAPPY mapping, 318, 323
 principle, 323
 Hessian Fly, 560
Hessian fly-response (Hfr) genes, 574
Heteranthelium piliferum, 38
 High density mapping
 difficulties, 165
 High-resolution mapping
 improving confidence interval for
 QTL, 219
 obtaining markers closer to gene, 219
 steps, 220
 studies, 221
 High-resolution melting analysis
 (HRM), 222
 High-throughput fingerprinting,
 325–326, 327
 High-throughput functional analysis, 565
 genes silenced/overexpressed, 567–568
 Transient-Induced-Gene-Silencing
 (TIGS) pipeline, 566
Hooded mutation, 599
Hordeum bogdanii, 39
Hordeum brevisubulatum, 39
Hordeum bulbosum, 39, 138
Hordeum capense, 39
Hordeum jubatum, 39
Hordeum spontaneum (Wild barley), 91
Hordeum vulgare (Barley), 3, 31, 39, 81–82,
 90, 127, 137, 138, 202–203
 accessions per taxon, 55–56
 chromosomal mutations, 146
 chromosome sizes, 138
 core collections, 56
 databases, 54
 distribution, 138
 diversity and domestication, 91
 forms, 203
 genepool, 138
 genetic stocks collections, 57
 genome, 138–139
 homologous chromosomes, 337
 institutions with accessions, 53, 54, 58
 molecular markers comparing wild to
 domesticated, 92
 single *versus* multiple origins, 92
 two-row and six-row genotypes, 91–92

 uses, 82
 varieties, 91
 wild, distribution of, 91
 wild *vs.* domesticated, 92
 Host-pathogen interaction, 569
 Hypersensitive cell death (HR), 561

I

Illumina genome analyzer, 730
 Insertional mutagenesis, 376
 Insertion Site-Based Polymorphisms (ISBP),
 264, 306, 400
In silico comparative analyses, 455
In situ hybridization, 121, 123, 146
 Intercellular secretome, 563
 International Agricultural Research Centres
 (IARCs), 33
 The International Barley Core Collection
 (BCC), 56
 International Barley Genome Sequencing
 Consortium (IBSC), 57, 105, 330,
 345, 408
 International Code for Botanical
 Nomenclature, 5
 International Plant Names Index (IPNI), 23
 International Triticeae Consortium, 24
 International Triticeae EST Cooperative
 (ITEC), 409
 International Triticeae Mapping Initiative
 (ITMI), 105
 International Wheat Genome Sequencing
 Consortium (IWGSC), 52, 105,
 331, 345, 408
 International Wheat Information System
 (IWIS), 49
 Inter-retrotransposon amplified
 polymorphism (I-RAP), 262
 Interval mapping, 217
 Introgression lines (ILs), 151
 Intron/exon boundaries, 469
 Isolation of deletion stocks, 123, 127

J

Jackknife procedure, 165
 JoinMap, 205, 214

K

Karyotypes, 142
Kengyilia grandiglumis, 39
Knotted1 gene, 599

L

- Leucine rich repeat (LRR), 343, 560
- Leymus angustus*, 39
- Leymus arenarius*, 40
- Leymus cinereus*, 40
- Leymus mollis*, 40
- Leymus racemosus*, 40
- Leymus triticoides*, 40
- Leymus* (Wildrye), 39–40
- Linkage analysis
 - programs, 205
- Linkage disequilibrium (LD), 512
- Long terminal repeat (LTR), 262
- Loss-of-function mutants, 371
- Low nutrients, tolerance to
 - nitrogen, 523–525
 - phosphorus, 525–526
- LTR_STRUC program, 418

M

- Macrocolinearity studies, 454
 - marker based, 454–455
 - sequence based, 455–458
- Maize, 723–724
- Mansfeld's Encyclopedia, 32, 35
- Map-based cloning, 337, 338, 563
 - application and problems of chromosome walking, 345–346
 - aspects of sequencing and identification, 347–348
 - genes isolated from wheat and barley, 338–343
 - genetic mapping, 343–344
 - list of genes cloned, 340–342
 - outlook, 353
 - physical mapping for, 345
 - problems caused by repetitive elements, 346–347
 - role of bioinformatics, 352–353
 - strategy to isolate gene of interest, 339
 - use and limits of model genomes, 348–350
 - comparison of *Lr34* locus from wheat, 349
 - validation of candidate genes, 350–352
 - gene density, 350
 - See also* positional cloning
- Map construction
 - dominant and codominant markers, 174
- Map curation
 - factors affecting quality of, 211–212
 - identifying map quality issues, 212
 - inspection of marker genotype data, 212
 - simulated chromosome with markers, 213

- MapManager QTX, 205
- Mapping algorithms, 168
- Marker order, 164
- Marker validation, 218
- Meiosis in wheat- and role of *Ph1*, early
 - stages of, 237–249
 - chromosome pairing loci, 241–242
 - exploitation of, 248–249
 - in presence and absence of *Ph1* locus, 242
 - chromosome sorting, 238–239
 - Ph1* locus, 242–248
 - polyploids, 240–241
 - recombination, 239–240
- Meiotic pairing analysis, 121, 122
- Microarrays, 268–269
- Microbe-associated-molecular-patterns (MAMPs), 569
- Microcolinearity
 - comparative sequencing, 461–462
 - inter and intra specific, 459
 - locus-based, 458–464
 - sequence rearrangement, violation of, 462–464
- Microdissection, 287–288
- MicroRNAs, 563
- Microsatellites, 258
 - See also* Simple Sequence Repeats (SSRs)
- Minimal-map-length criterion, 164
- Minimal Tiling Path (MTP), 324
- Missouri Botanical Garden, 23
- Mla*-mediated barley-powdery mildew interactions, 569
- Mlo*-resistance gene, 571
- Model genomes
 - use and limits of, 348–350
- Molecular markers, 83, 204, 217
 - AFLP markers, 260–261
 - PCR based, 344
 - repeat-based, 261–265
 - types of, 263
 - RFLP clones, 256–258
 - SSR markers, 258–260
 - technology, 202
 - use of, 83
- Molecular plant-microbe interactions, 569
- Molecule profiling approaches, 562–563
- Monosomics, 123
- mRNA profiling, 562, 574
- Multilocus consensus genetic mapping (MCGM), 172
 - problems, 173

- Multilocus mapping
 - criteria of comparisons, 171
 - objectives, 166–167
 - problems, 164
 - stages in procedure, 170
 - usage of, 163
- Multi-parent Advanced Generation
 - Intercross (MAGIC), 208
- Multiple interval mapping (MIM), 175, 184–187
- Multiple QTL mapping, 217
- Multiple-trait (MT) analysis
 - application of, 177
- Multipoint, 205
- Mutagenesis, 360
- Mutagens, 360
- Mutation
 - detection, 364–365
 - in DNA, 362
 - frequency, 360–362
- Mycosphaerella graminicola* (*Septoria Tritici* blotch), 576
 - genomic sequence, 576
- N**
 - N-banding, 138, 144, 150
 - Near Isogenic Lines (NILs), 344
 - Negative crossover interference, 168, 169
 - Neolithic Revolution, 82
 - Nomenclatural problems of taxonomic changes, 7
 - Non-homologous chromosomes
 - random segregation of, 167
 - Nonhost resistance functions, 561
 - NS-orders, 174
 - Nucleotide binding site (NBS), 343, 560
 - Nud* gene, 94
 - Nulli-tetrasomics, 123
 - N-use efficiency (NUE), 523
- O**
 - Oasis Theory, 82
 - Oligonucleotide arrays, 269
 - “-omics” platforms, 513
 - Online databases, 74
 - Optical mapping, 210
 - Optimal consensus order, 173
 - Ordered-clone based physical mapping
 - chromosome walking, 327
 - Ordered-marker based physical mapping
 - cytogenetic stocks and chromosome-microdissection, 319–321
 - Fluorescence In Situ Hybridization (FISH), 321–322
 - radiation hybrid mapping (RH) and HAPPY mapping, 322–323
 - Osmotic adjustment (OA), 517
- P**
 - Parameters of chlorophyll fluorescence kinetics (PCFKs), 520
 - Particle bombardment, 436
 - Pascopyrum smithii*, 40
 - Pathogen Associated Molecular Patterns (PAMP), 561, 569
 - Pathogenesisrelated (PR) genes, 561
 - Pathogen genomics, 574–579
 - PCR analysis, 146
 - PCR-based molecular markers, 258
 - PCR-based techniques, 205
 - PEN1* and *Ror2* genes, 571
 - Pgt* TTKS, 560
 - PhenoMap, 214
 - Photoperiod insensitive phenotype, 598, 601
 - Photoperiod (*Ppd* genes), 102–103, 593, 600–601
 - alleles, 598
 - Ppd-1*, 598
 - Ph1* (Pairing homoeologous 1), 241–248
 - Physical mapping, 318, 441
 - barley chromosome 5H, 320
 - deletion panel of barley chromosome, 320
 - of diploid Triticeae genomes
 - Aegilops Tauschii*, 329–330
 - Barley (*Hordeum vulgare*), 330
 - generating
 - optical mapping, 327–328
 - ordered-clone based, 324–327
 - ordered-marker based, 318–324
 - of polyploid Triticeae genomes
 - Bread wheat (*Triticum aestivum*), 331
 - of Sec-1 locus, 211
 - ongoing projects, 329
 - ordered genomic clones, construction of, 324
 - Triticeae genomes
 - diploid, 329
 - polyploid, 331
 - whole genome, building, 324
 - “The phylogeographical basis for plant breeding”, 82

Pina and *Pinb* gene, 100
 Plant chromosomes, isolation of, 291
 Plant diseases, 559
 Plant domestication
 understanding, 81
 using molecular biology, 105
 Plant genome projects, 724
 Plant-insect interactions, 573
 Plant ontology, 370–371
 Plant-resistance (R) genes, 560
 Poaceae family, 451–452
 Polymerase Chain Reaction (PCR), 344
 Polymorphism detection, 205
 Polyploids, 240
 Pooideae family, 452
 Positional cloning, 327, 512
 challenges, 338
 Positive crossover interference, 168
 See also Negative crossover interference
Ppd mutation, 600–601
 Principle of Priority, 7
 Project Aligned Related Sequences and
 Evaluate SNPs (PARSESNP), 362
 Protein coding genes
 functional annotation of, 407
 Protein Quantity Loci (PQLs), 514
Psathyrostachys juncea, 40
 Pseudo-linkage, 167, 168
Pseudo-Response Regulator (PRR)
 genes, 598
Pseudoroegneria spicata, 7, 41
Puccinia graminis (stem rust), 575–576
 genomic sequence, 576
 host range, 575–576
Pyrenophora teres (net blotch), 560

Q

Q mutation, 599–600
 QTL × environment interactions (QEI), 175,
 181, 183–184, 534–535
 QTL mapping, 163, 181, 522, 564
 expression-, 174–175
 for loci controlling flowering
 time, 104
 multiple approaches, 174–175
 Multiple Environment-Multilocus
 (ME-MIM), 174
 Multiple Trait-Multilocus
 (MT-MIM), 178
 QTL detection power, 174–175,
 178, 186
 ‘QTL pipeline,’ 224

Quantitative trait loci (QTL), 163, 217,
 467, 481
 allele substitution, 178
 analysis, 178
 cloning, 512
 detection methods, 217
 ordering of loci within linkage
 groups, 219
 Quasi-linkage, 167

R

Radiation Hybrid Mapping
 (RH), 318
 Random amplified microsatellite
 polymorphisms (RAPs), 145
 Random Amplified Polymorphic DNAs
 (RAPDs), 344
 Recombinant inbred lines (RILs), 517
 RECORD, 205
Reduced Height-1 (Rht-1)
 cloning of, 595
 Restriction Fragment Length
 Polymorphisms (RFLPs), 145,
 243, 256, 343
 markers, 205, 256, 454–455
 Retrotransposon-Based Insertional
 Polymorphism (RBIP), 264
 Retrotransposon markers, 145–146
 RETrotransposon-Microsatellite Amplified
 Polymorphism (REMAP), 262
Rht-8, 100
Rht-B1b and *Rht-D1b* genes, 99
Rht-1 gene, 99
Rht genes, 100
 Rice, 720–722
 Rice full length cDNAs, 469
 RNA interference (RNAi), 374–375
 based gene silencing, 352
 techniques, 565
 RNA isolation, 563
 Roche-454 GSFLX, 729–730
 Royal Botanic Gardens, Kew, 24

S

Scientific names in Triticeae, 3–24
 interaction of taxonomy and
 nomenclature
 multiple names at generic level, 7–8
 multiple names at species level and
 below, 9–12
 problems with generic changes, 8–9

- multiple names, 4–5
 - impact of new technologies on taxonomy, 5–7
 - nomenclatural web sites, 23–24
 - taxonomic treatment, 4–7
 - genera, 14–22
 - species, 22–23
 - Secale cereale* (rye), 3, 31, 41, 81, 138, 202, 203
 - accessions per species, 63
 - chromosomes, 128
 - cultivated, 93–94
 - databases, 63
 - genetic stocks collections, 63
 - groups of taxa, 93
 - institutions, 62
 - uses, 82
 - Secale × derzhavinii*, 41
 - Secale strictum*, 41
 - Secale vavilovii*, 41, 93
 - Seed dormancy, 102
 - Selective DNA pooling (SDP) analysis, 189
 - Sequence assembly, 417
 - Sequence contig, 417
 - Sequence-Specific Amplified Polymorphism (SSAP), 262, 400
 - Sequence Tagged Sites (STS), 344
 - Simple interval mapping (SIM), 184–188
 - Simple Sequence Repeats (SSRs), 258–260, 344
 - based markers, 145–146
 - Single-cell assay, 565–566
 - Single-cell transcriptomes, 563
 - Single chromosome addition/substitution lines, 323
 - Single marker analysis, 217
 - Single Nucleotide Polymorphisms (SNPs), 205, 266–268
 - markers, 344
 - Single strand Conformational Polymorphism (SSCP), 469
 - SNPs, *see* Single Nucleotide Polymorphisms (SNPs)
 - S-orders, 174
 - Sorting Intolerant from Tolerant (SIFT), 362
 - Species relationships, assessing, 142–145
 - SSCP-SNP marker technique, 469
 - SSR (microsatellites), 150
 - Stable transformation, 371
 - Stagonospora nodorum* (*Stagonospora Nodorum* Blotch)
 - genomic sequence, 577
 - wheat pathogen, 577
 - Stem rust, 560
 - Synapsis, 237
 - Synaptonemal complex (SC), 237
 - Systemic acquired resistance (SAR), 561
 - System-wide Information Network for Genetic Resources (SINGER), 33
- T**
- Tagged microarray marker (TAM), 264
 - Tandem repeated sequences, structure of
 - Triticeae genome, 687–688
 - centromeric regions, 688–690
 - interstitial sites, 692–693
 - locations and translocations, 688
 - rRNA genes, 693–695
 - telomeric region, 690–692
 - Targeting Induced Local Lesions in Genomes (TILLING), 565
 - Target-Primed Reverse Transcription (TPRT), 393
 - Target site duplication (TSD), 416
 - “Taxonomic Nomenclature Checker” (TNC-GRIN)
 - stepwise process, 35
 - Taxonomic treatment, 4
 - morphologically distinct genera, 13
 - TEnest program, 419
 - Thinopyrum elongatum*, 42
 - Thinopyrum intermedium*, 42
 - Thinopyrum junceiforme*, 42
 - Thinopyrum junceum*, 42
 - Thinopyrum ponticum*, 42
 - Thinopyrum* (Wheatgrass), 42
 - Tillering and branching mutations, 101
 - TILLING, 360, 515
 - computational tools, 362–363
 - development of resource, 364
 - Tin3* gene, 101
 - Toolbox for Triticeae genomics
 - BAC libraries, 269–272
 - EST sequences and microarrays, 268–269
 - molecular markers
 - AFLP markers, 260–261
 - DARt markers, 265–266
 - DNA clones to detect RFLPs, 257
 - repeat-based markers, 261–265
 - RFLP clones, 256–258
 - SNP arrays, 266–268
 - SSR markers, 258–260
 - outlook, 272–273

- Transcript-based cloning, 563–564
 - fast neutron mutants, 564
 - Transcriptome profiling, 514–515
 - Transgenesis
 - in gene function, 376
 - resistance gene functions, 377
 - Transient-induced-genesilencing (TIGS), 565
 - Transposable Elements (TEs), 261–262, 460
 - classification, 389
 - databases, 419–420
 - as molecular markers, 400
 - recombinational loss, 399
 - structures
 - Helitron, 395–396
 - LINE elements, 391, 393
 - LTR retrotransposons, 391
 - SINE elements, 393–394
 - TIR elements, 395
 - Transposon Display (TD), 400
 - Traveling salesperson problem (TSP), 164
 - Trial-and-error approach, 170
 - Trithothecene
 - classes of genes, 571
 - mycotoxin accumulation, 570
 - Triticale (\times *Triticosecale*), 203
 - Triticeae chromosomes
 - centromere-telomere axis of, 699–700
 - cause of correlation between gene density and recombination rate, 701–702
 - evolutionary significance of repeated DNA, 702–703
 - variation in gene density along chromosomes, 700–701
 - rDNA in, 688
 - Triticeae genomes, 387
 - Triticeae genomic sequences
 - annotating, 411
 - Triticeae repeat database (TREP), 420
 - \times *Triticosecale* (Triticale), 3, 31, 42
 - databases, 57
 - genetic stocks collections, 59
 - institutions with accessions, 58–59, 60
 - Triticum aestivum* (bread wheat), 3, 31, 42, 89–90, 138, 202
 - free-threshing, 94
 - Triticum aethiopicum*, 42
 - Triticum araraticum*, 42, 88, 89
 - Triticum boeoticum*, 42, 85
 - Triticum dicoccoides*, 43, 88
 - phylogenetic analysis, 88
 - Triticum dicoccon*, 43
 - Triticum dicoccum*, 88, 89
 - Triticum durum*, 3, 31–32, 43
 - free-threshing, 94
 - Triticum ispahanicum*, 43
 - Triticum karamyshevii*, 43
 - Triticum macha*, 43
 - Triticum monococcum* (einkorn wheat), 9–12, 43, 85
 - names of close relatives
 - of, 10
 - Triticum persicum*, 43
 - Triticum polonicum*, 43
 - Triticum sinkajae*, 43
 - Triticum spelta*, 43, 90
 - Triticum sphaerococcum*, 43
 - Triticum timopheevii*, 43
 - Triticum turanicum*, 43–44
 - Triticum turgidum*, 44, 138
 - Triticum urartu*, 44, 85, 87
 - Triticum vavilovii*, 44
 - Triticum* (wheat)
 - homoeologous genomes, 337
 - uses, 82
 - Triticum zhukovskyi*, 44
 - \times *Tritordeum*, 44
- U**
- Uniparental chromosome
 - elimination, 149
 - United States Department of Agriculture (USDA-ARS), 33, 35
 - USDA National Plant Germplasm System (NPGS), 434
 - USDA online information, 35
- V**
- Vernalization pathway
 - VRN1* and *VRN2* genes controlling, 103–104
 - Vernalization (*Vrn* genes), 593
 - Vrn-1*, 597
 - Vrn-2*, 596–597
 - Vrn-3*, 597–598
 - Virus-induced gene silencing (VIGS), 351, 366, 568
 - Vrs1* gene, 599

W

WCM approach, 186

Wheat and barley genome sequencing,
713–716

applied biosystems SOLiD (sequencing
by oligo ligation and detection),
730–731

HeliScope, Helicos, 731

impact on Triticeae genome
sequencing, 731–733

contiguous genomic DNA sequences,
727–728

current status of Triticeae genome
sequencing, 725

EST sequencing, 725–726

GSS, 726–728

future of Triticeae genome sequencing,
733–735

history of sequencing in higher plants,
716

first plant genome model –
Arabidopsis thaliana, 720

first tree genome – poplar genome
sequence, 722–723

higher plant genome sequencing
projects, completed, draft, and
ongoing, 717–719

rice – economically important plant
genome, 720–722

illumina genome analyzer, 730

next generation sequencing (NGS)
technologies, 728–733

outlook, 735

Roche-454 GSFLX, 729–730

two grapevine sequences, 723

maize, 723–724

plant genome projects, 724

Wheat and rye genomes, cytogenetic analysis
of, 121–131

future prospects, 130–131

phases of cytogenetic research, 122–124

rye anchor karyotype, 128–130

C-banding, 128–129

comparative mapping, 129

deletion chromosomes, 130

idiogram representation, 128

wheat anchor karyotype, 124–126

idiogram representation, 125

wheat chromosome differentiation,
126–128

chromosome behavior, 127

Wheat evolution and domestication, 84–90

diploid, 85–88

dispersed-specific model, 87

genetic relationships, 85

hexaploid (bread wheat), 89–90

tetraploid, 88–89

Wheat flour

groups of proteins, 622

Wheat GeneChip assays, 562–563

Wheat stem rust (Ug99), 560

Wheat Strategy, 49

Wheat streak mosaic virus (WSMV), 573

Wheat (*Triticum* spp.)

levels of ploidy, 202–203

Whole genome amplification, 323

Whole-genome linkage map, 204–205

Whole-genome shotgun (WGS), 574

Wild and domesticated species

of Neolithic agriculture, 84

World Economic Plants Report, 33

“World Information and Early Warning
System” (WIEWS), 33

Y

Yield losses

caused by abiotic stress, 515

Z

Zn deficiency, 529